

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 03:24:37 ; Search time 216 Seconds  
(without alignments)  
6219.278 Million cell updates/sec

Title: US-09-787-844-1  
Perfect score: 1081  
Sequence: 1 aggagcagagggggcgta.....acattccagtgtgctgc 1081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1055.8	97.7	1100	9	US-09-905-291A-256
2	1055.8	97.7	1100	9	Sequence 256, App
3	1055.8	97.7	1100	9	Sequence 256, App
4	1055.8	97.7	1100	9	Sequence 256, App
5	1055.8	97.7	1100	9	Sequence 256, App
6	1055.8	97.7	1100	9	Sequence 256, App
7	1055.8	97.7	1100	9	Sequence 256, App
8	1055.8	97.7	1100	9	Sequence 256, App
9	1055.8	97.7	1100	9	Sequence 256, App
10	1055.8	97.7	1100	9	Sequence 256, App
11	1055.8	97.7	1100	9	Sequence 256, App
12	1055.8	97.7	1100	9	Sequence 256, App
13	1055.8	97.7	1100	9	Sequence 256, App
14	1055.8	97.7	1100	9	Sequence 256, App
15	1055.8	97.7	1100	9	Sequence 256, App
16	1055.8	97.7	1100	9	Sequence 256, App
17	1055.8	97.7	1100	9	Sequence 256, App
18	1055.8	97.7	1100	9	Sequence 256, App
19	1055.8	97.7	1100	9	Sequence 256, App

20	1055.8	97.7	1100	9	US-09-904-119-256
21	1055.8	97.7	1100	9	US-09-904-956-256
22	1055.8	97.7	1100	9	US-09-907-794-256
23	1055.8	97.7	1100	9	US-09-902-692-256
24	1055.8	97.7	1100	9	US-09-903-520-256
25	1055.8	97.7	1100	9	US-09-903-943-256
26	1055.8	97.7	1100	9	US-09-904-462-256
27	1055.8	97.7	1100	9	US-09-905-056-256
28	1055.8	97.7	1100	9	US-09-907-925-256
29	1055.8	97.7	1100	9	US-09-904-533-256
30	1055.8	97.7	1100	9	US-09-905-381-256
31	1055.8	97.7	1100	9	US-09-905-084-256
32	1055.8	97.7	1100	9	US-09-905-088-256
33	1055.8	97.7	1100	9	US-09-907-575-256
34	1055.8	97.7	1100	9	US-09-902-759-256
35	1055.8	97.7	1100	9	US-09-905-075-256
36	1055.8	97.7	1100	9	US-09-902-634-256
37	1055.8	97.7	1100	9	US-09-902-713-256
38	1055.8	97.7	1100	9	US-09-907-979-256
39	1055.8	97.7	1100	10	US-09-909-320-256
40	1055.8	97.7	1100	10	US-09-909-088-256
41	1054.8	97.6	1077	10	US-09-885-441-20
42	1045.8	96.7	1082	10	US-09-885-441-19
43	493.4	45.6	529	10	US-09-833-790-374
C 44	465.4	43.1	575	9	US-10-015-219-687
C 45	465.4	43.1	575	10	US-09-777-564-687

ALIGNMENTS

RESULT 1  
US-09-905-291A-256  
; Sequence 256, Application US/09905291A  
; Patent No. US20020160374A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secrated and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/905.291A  
; CURRENT FILING DATE: 2001-07-12  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222

;; PRIOR FILING DATE: 1999-07-28  
;; PRIOR APPLICATION NUMBER: PCT/US99/20594  
;; PRIOR FILING DATE: 1999-09-08  
;; PRIOR APPLICATION NUMBER: PCT/US99/20944  
;; PRIOR FILING DATE: 1999-09-13  
;; PRIOR APPLICATION NUMBER: PCT/US99/21090  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/21547  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/23089  
;; PRIOR FILING DATE: 1999-10-05  
;; PRIOR APPLICATION NUMBER: PCT/US99/28214  
;; PRIOR FILING DATE: 1999-11-29  
;; PRIOR APPLICATION NUMBER: PCT/US99/28313  
;; PRIOR FILING DATE: 1999-11-30  
;; PRIOR APPLICATION NUMBER: PCT/US99/28564  
;; PRIOR FILING DATE: 1999-12-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/28565  
;; PRIOR FILING DATE: 1999-12-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/30095  
;; PRIOR FILING DATE: 1999-12-16  
;; PRIOR APPLICATION NUMBER: PCT/US99/30911  
;; PRIOR FILING DATE: 1999-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US99/30999  
;; PRIOR FILING DATE: 1999-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US00/00219  
;; PRIOR FILING DATE: 2000-01-05  
;; NUMBER OF SEQ ID NOS: 423  
;; SEQ ID NO 256  
;; LENGTH: 1100  
;; TYPE: DNA  
;; ORGANISM: Homo Saplen  
US-09-787-844-1

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 21 GCGCGGGAGAGAGGCCATGGCGCGCGGGGGCTGCTGCTGGCGCTGCTGCTGGC 80  
DB 1 GCGCGGGAGAGAGAGGCCATGGCGCGCGGGGGCTGCTGCTGGCGCTGCTGCTGGC 60  
QY 81 TCGGGCTGGACTCAGAGAGCGGAGTCGAGGAGCGGGCCCTTATCAGGACCATGCGG 140  
DB 61 TCGGGCTGGACTCAGAGAGCGGAGTCGAGGAGCGGGCCCTTATCAGGACCATGCGG 120  
QY 141 CCAGCGGTCATCAGCTCGCGATCGTGGGATTCGCCACGATGCGGAGTGCAGCTGCTCAGCCA 260  
DB 121 CCAGCGGTCATCAGCTCGCGATCGTGGGATTCGCCACGATGCGGAGTGCAGCTGCTCAGCCA 240  
QY 201 GTGGCAGGGAGCGCTCGCGCTGCTGGGATTCGCCACGATGCGGAGTGCAGCTGCTCAGCCA 260  
DB 181 GTGGCAGGGAGCGCTCGCGCTGCTGGGATTCGCCACGATGCGGAGTGCAGCTGCTCAGCCA 240  
QY 261 CGGCTGGGACATCAGCGGCGGACATGCTTTGAAACCTATAGTGACCTTAGTCCCTC 320  
DB 241 CGGCTGGGACATCAGCGGCGGACATGCTTTGAAACCTATAGTGACCTTAGTCCCTC 300  
QY 321 CGGCTGGGATGCTCAGCTGGGCGGACGTCGCTTCATGCCATCCTTCTGAGGCGCTGCAGGC 380  
DB 301 CGGCTGGGATGCTCAGCTGGGCGGACGTCGCTTCATGCCATCCTTCTGAGGCGCTGCAGGC 360  
QY 381 CTACTACACCCGCTTACTTCTGATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 440  
DB 361 CTACTACACCCGCTTACTTCTGATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 420  
QY 441 ACCCTATGACATGCTTGGTGAAGCTGTGTGCACCTGTACCTACACTAACACATCCA 500  
DB 421 ACCCTATGACATGCTTGGTGAAGCTGTGTGCACCTGTACCTACACTAACACATCCA 480  
QY 501 GCCCATCTGTCTCCAGGCGCTCCACATTTAGTTTGGAGAACCGGACGAGACTGCTGGGTGAC 560  
DB 481 GCCCATCTGTCTCCAGGCGCTCCACATTTAGTTTGGAGAACCGGACGAGACTGCTGGGTGAC 540

QY 561 TGGCTGGGGGTACATCAAGAGAGTGGGACTGCCATCTCCACACACCCCTCCAGGAAGT 620  
DB 541 TGGCTGGGGGTACATCAAGAGAGTGGGACTGCCATCTCCACACACCCCTCCAGGAAGT 600  
QY 621 TCAGTGCCTCATATAAACAACCTCTATGTGCAACACACCTCTCTCTCAAGTACAGTTTCG 680  
DB 601 TCAGTGCCTCATATAAACAACCTCTATGTGCAACACACCTCTCTCTCAAGTACAGTTTCG 660  
QY 681 CAAGGACATCTTTGGAGACATGGTTTGTCTGGCAATGCCCAAGCGGGAAGGATGCCCTG 740  
DB 661 CAAGGACATCTTTGGAGACATGGTTTGTCTGGCAATGCCCAAGCGGGAAGGATGCCCTG 720  
QY 741 CTTGGTGACTCAGGTGGACCCCTTGGGCTGTAAACAAGATGGACTGTGTATCAGATTGG 800  
DB 721 CTTGGTGACTCAGGTGGACCCCTTGGGCTGTAAACAAGATGGACTGTGTATCAGATTGG 780  
QY 801 AGTGTGAGCTGGGAGTGGGCTGTGTGCGGCCCAATGCGGCCCGGCTGTCTACACCAATAT 860  
DB 781 AGTGTGAGCTGGGAGTGGGCTGTGTGCGGCCCAATGCGGCCCGGCTGTCTACACCAATAT 840  
QY 861 CAGCCACCATCTTTGAGTGGATCCAGAAAGCTGATGGCCAGAGTGGCATGTCCAGGCCAGA 920  
DB 841 CAGCCACCATCTTTGAGTGGATCCAGAAAGCTGATGGCCAGAGTGGCATGTCCAGGCCAGA 900  
QY 921 CCCCTCTGGCCACTACTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 980  
DB 901 CCCCTCTGGCCACTACTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960  
QY 981 CTGAGCTACTCTGAGCCCATGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1040  
DB 961 CTGAGCTACTCTGAGCCCATGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1020  
QY 1041 CTCTCTTTGTTGTTAATAACACATCTCCAGTTGATGCT 1079  
DB 1021 CTCTCTTTGTTGTTAATAACACATCTCCAGTTGATGCT 1059

## RESULT 2

US-09-787-844-1  
; Sequence 256, Application US/09902853  
; Publication No. US20020192659A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/902,853  
; CURRENT FILING DATE: 2001-07-10

;; PRIOR APPLICATION NUMBER: US/09/665,350  
;; PRIOR FILING DATE: 2000-09-18  
;; PRIOR APPLICATION NUMBER: US 60/143,048  
;; PRIOR FILING DATE: 1999-07-07  
;; PRIOR APPLICATION NUMBER: US 60/145,698  
;; PRIOR FILING DATE: 1999-07-26  
;; PRIOR APPLICATION NUMBER: US 60/146,222  
;; PRIOR FILING DATE: 1999-07-28  
;; PRIOR APPLICATION NUMBER: PCT/US99/20594  
;; PRIOR FILING DATE: 1999-09-08  
;; PRIOR APPLICATION NUMBER: PCT/US99/20944  
;; PRIOR FILING DATE: 1999-09-13  
;; PRIOR APPLICATION NUMBER: PCT/US99/21090  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/21547  
;; PRIOR FILING DATE: 1999-09-15  
;; PRIOR APPLICATION NUMBER: PCT/US99/23089  
;; PRIOR FILING DATE: 1999-10-05  
;; PRIOR APPLICATION NUMBER: PCT/US99/28214  
;; PRIOR FILING DATE: 1999-11-29  
;; PRIOR APPLICATION NUMBER: PCT/US99/28313  
;; PRIOR FILING DATE: 1999-11-30  
;; PRIOR APPLICATION NUMBER: PCT/US99/28564  
;; PRIOR FILING DATE: 1999-12-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/28565  
;; PRIOR FILING DATE: 1999-12-02  
;; PRIOR APPLICATION NUMBER: PCT/US99/30095  
;; PRIOR FILING DATE: 1999-12-16  
;; PRIOR APPLICATION NUMBER: PCT/US99/30911  
;; PRIOR FILING DATE: 1999-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US99/30999  
;; PRIOR FILING DATE: 1999-12-20  
;; PRIOR APPLICATION NUMBER: PCT/US00/00219  
;; PRIOR FILING DATE: 2000-01-05  
;; NUMBER OF SEQ ID NOS: 423  
;; SEQ ID NO 256  
;; LENGTH: 1100  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-09-902-853-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 GGC CGCGG GAGAGGAGGCATCGGCGCGCGCGCTGCTGCTGGCGCTGCTGCTGCG 80  
DB 1 GGC CGCGG GAGAGGAGGCATCGGCGCGCGCGCTGCTGCTGGCGCTGCTGCTGCG 60  
QY 81 TCGGCGTGGACTCAGGAAGCCGAGTGCAGGAGCGCGCGCTTATCAGGACCATGCGG 140  
DB 61 TCGGCGTGGACTCAGGAAGCCGAGTGCAGGAGCGCGCGCTTATCAGGACCATGCGG 120  
QY 141 CGGACGGGTATCAGCTGCGGATCGTGGTGAGAGAGCGCGGAGTGGCGGTGGCC 200  
DB 121 CGGACGGGTATCAGCTGCGGATCGTGGTGAGAGAGCGCGGAGTGGCGGTGGCC 180  
QY 201 GTGGCAGGGGAGCGCTGCGCTGCGGATTCGCCACGATGCGGAGTGCCTGCAGCA 260  
DB 181 GTGGCAGGGGAGCGCTGCGCTGCGGATTCGCCACGATGCGGAGTGCCTGCAGCA 240  
QY 261 CGGCTGGGCACATCAGCGCGGCGCACTGCTTTGAAACCTATAGTACCTTAGTACCTTC 320  
DB 241 CGGCTGGGCACATCAGCGCGGCGCACTGCTTTGAAACCTATAGTACCTTAGTACCTTC 300  
QY 321 CGGCTGGATGGTCCAGTGTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCGCTGAGCG 380  
DB 301 CGGCTGGATGGTCCAGTGTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCGCTGAGCG 360  
QY 381 CTACTACACCGCTTACTTCGATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 440  
DB 361 CTACTACACCGCTTACTTCGATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 420

QY 441 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTCACTACACTAAACACATCCA 500  
DB 421 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTCACTACACTAAACACATCCA 480  
QY 501 GCCCATCTCTCTCCAGGCGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC 560  
DB 481 GCCCATCTCTCTCCAGGCGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC 540  
QY 561 TGGCTGGGGGTACATCAAGAGAGTGAAGGACCTGCCATCTCCCCACACCTCCAGGAAGT 620  
DB 541 TGGCTGGGGGTACATCAAGAGAGTGAAGGACCTGCCATCTCCCCACACCTCCAGGAAGT 600  
QY 621 TCAGGTGCGCATCATATAACAACACTTATGTGCAACCACTCTTCTCCTCAAGTACAGTTTCG 680  
DB 601 TCAGGTGCGCATCATATAACAACACTTATGTGCAACCACTCTTCTCCTCAAGTACAGTTTCG 660  
QY 681 CAAGGACATCTTTGGAGACATGGTTTGTGTCGCAATGCCAATGCCAAGCGGGAAGATGCTG 740  
DB 661 CAAGGACATCTTTGGAGACATGGTTTGTGTCGCAATGCCAATGCCAAGCGGGAAGATGCTG 720  
QY 741 CTTGCTGACTCAGGTGGGACCTTGGCCTGTAAACAAGATGGACTGTGTATCAGATTGG 800  
DB 721 CTTGCTGACTCAGGTGGGACCTTGGCCTGTAAACAAGATGGACTGTGTATCAGATTGG 780  
QY 801 AGTCTGTAGCTGGGAGTGGGCTGTGTCGCGCCCAATCGCGCGCTGTCTACACCAATAT 860  
DB 781 AGTCTGTAGCTGGGAGTGGGCTGTGTCGCGCCCAATCGCGCGCTGTCTACACCAATAT 840  
QY 861 CAGCCACCATCTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGCCAGCCAGA 920  
DB 841 CAGCCACCATCTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGCCAGCCAGA 900  
QY 921 CCCCTCTCGCCACTACTCTTTTCCCTCTCTCTGGCTCTCCCACTCTGGGGCGGT 980  
DB 901 CCCCTCTCGCCACTACTCTTTTCCCTCTCTCTGGCTCTCCCACTCTGGGGCGGT 960  
QY 981 CTGAGCCTACCTGAGCCCATGAGCCTGGGGCCACTGCCAAGTGCAGGCCCTGTTCTCTT 1040  
DB 961 CTGAGCCTACCTGAGCCCATGAGCCTGGGGCCACTGCCAAGTGCAGGCCCTGTTCTCTT 1020  
QY 1041 CTGCTTGTGTTGTTAATAACACATTCAGTTGATGCCT 1079  
DB 1021 CTGCTTGTGTTGTTAATAACACATTCAGTTGATGCCT 1059

RESULT 3  
US-09-907-824-256  
; Sequence 256, Application US/09907824  
; Publication No. US20020197671A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/907,824  
CURRENT FILING DATE: 2001-07-17  
PRIOR APPLICATION NUMBER: 09/665,350  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 256  
LENGTH: 1100  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-09-907-824-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	21	GGCCGCGGAGAGGAGCGCATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGC	80
Db	1	GGCCGCGGAGAGGAGCGCATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGC	60
QY	81	TCGGGCTGACTCAGGAAGCGGAGTCGCGAGGCGCGCGCTTATCAGGACCATGGCG	140
Db	61	TCGGGCTGAGCTCAGGAAGCGGAGTCGCGAGGCGCGCGCTTATCAGGACCATGGCG	120
QY	141	CCGACGGGTATCAGCTCGCGCATGGTGGTGAGAGGAGCGCGAAGTGGCGGTGGCG	200
Db	121	CCGACGGGTATCAGCTCGCGCATGGTGGTGAGAGGAGCGCGAAGTGGCGGTGGCG	180
QY	201	GTGGCAGGAGGAGCGCTGCGGCTTGGGATTCACAGTATCGGAGTGAGCGCTCTCAGCA	260
Db	181	GTGGCAGGAGGAGCGCTGCGGCTTGGGATTCACAGTATCGGAGTGAGCGCTCTCAGCA	240
QY	261	CGGCTGGGACATCAGCGGGCGCATGCTTTGAAACCTATAGTACCTTAGTGATCCCTC	320
Db	241	CGGCTGGGACATCAGCGGGCGCATGCTTTGAAACCTATAGTACCTTAGTGATCCCTC	300

QY	321	CGGGTGGATGGTCCAGTTGGCCAGCTGACTTCCATGCCATCTTCTTGAGGCTTGAGCG	380
Db	301	CGGGTGGATGGTCCAGTTGGCCAGCTGACTTCCATGCCATCTTCTTGAGGCTTGAGCG	360
QY	381	CTACTACACCGGTACTTTCGATATCGAATATCTATCTGAGCCCTCTGCTACCTGGGAATTC	440
Db	361	CTACTACACCGGTACTTTCGATATCGAATATCTATCTGAGCCCTCTGCTACCTGGGAATTC	420
QY	441	ACCCCTATGACATTTGGCTTTGGTGAAGCTGCTGCACCTCTGACCTACACTAAACACATCCA	500
Db	421	ACCCCTATGACATTTGGCTTTGGTGAAGCTGCTGCACCTCTGACCTACACTAAACACATCCA	480
QY	501	GCCCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC	560
Db	481	GCCCATCTGCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC	540
QY	561	TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCTCCAGGAAGT	620
Db	541	TGGCTGGGGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACACCTCCAGGAAGT	600
QY	621	TCAGGTGCCATCATAAACAACTCTATGTGCAACACCTCTTCTCAAGTACAGTTTCGG	680
Db	601	TCAGGTGCCATCATAAACAACTCTATGTGCAACACCTCTTCTCAAGTACAGTTTCGG	660
QY	581	CAAGGACATCTTTGGAGACATGGTTTGTGTGGCAATGCCCAAGCGGGAAGGATGCTG	740
Db	661	CAAGGACATCTTTGGAGACATGGTTTGTGTGGCAATGCCCAAGCGGGAAGGATGCTG	720
QY	741	CTTCGGTGACTCAGGTGGACCTTGGCTTGTAAACAAGTGGATGCTGATCAGATTGG	800
Db	721	CTTCGGTGACTCAGGTGGACCTTGGCTTGTAAACAAGTGGATGCTGATCAGATTGG	780
QY	801	AGTCGTGAGCTGGGGAGTGGCTGTGTGCGGCCCAATCGGCCGCTGCTACACCAATAT	860
Db	781	AGTCGTGAGCTGGGGAGTGGCTGTGTGCGGCCCAATCGGCCGCTGCTACACCAATAT	840
QY	861	CAGCCACCATTTGAGTGGATCCAGAACTGATGGCCAGAGTGGCATGTCCCAGCCAGA	920
Db	841	CAGCCACCATTTGAGTGGATCCAGAACTGATGGCCAGAGTGGCATGTCCCAGCCAGA	900
QY	921	CCCTCTCTGGCCACTACTCTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	980
Db	901	CCCTCTCTGGCCACTACTCTTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	960
QY	981	CTGAGCTTACTGAGCCCATGAGCCTGGGGCCACTGCGCAAGTCAGGCCCTGGTCTCTT	1040
Db	961	CTGAGCTTACTGAGCCCATGAGCCTGGGGCCACTGCGCAAGTCAGGCCCTGGTCTCTT	1020
QY	1041	CTGCTCTTTGGTAAATAACATTTCCAGTTGATGCTT 1079	
Db	1021	CTGCTCTTTGGTAAATAACATTTCCAGTTGATGCTT 1059	

## RESULT 4

US-09-907-841-256  
; Sequence 256, Application US/09907841  
; Publication No. US20020198366A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.















Db 841 CAGCCACCACTTGTAGTGGATCCAGAGCTGATGGCCAGAGTGCGATGTCCAGCCAGA 900  
Qy 921 CCCCTCCGCGCACTACTCTTCCCTCTCTCTGGGCTCTCCACACTCCCTGGGCGCGT 980  
Db 901 CCCCTCCGCGCACTACTCTTCCCTCTCTCTGGGCTCTCCACACTCCCTGGGCGCGT 960  
Qy 981 CTGAGCTTACCTGAGCCCATGAGCTGGGCGCACTGCAAGTCAAGGCGCTGTCTCTT 1040  
Db 961 CTGAGCTTACCTGAGCCCATGAGCTGGGCGCACTGCAAGTCAAGGCGCTGTCTCTT 1020  
Qy 1041 CTGCTCTTGGTAAATAACACATTCAGTTGATGCCT 1079  
Db 1021 CTGCTCTTGGTAAATAACACATTCAGTTGATGCCT 1059

RESULT 10

US-09-904-820-256

; Sequence 256, Application US/09904820  
; Publication No. US20030036094A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Flivaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/904,820  
; CURRENT FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 09/665,350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 256  
; LENGTH: 1100  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-904-820-256

Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 21 GGCCCGGAGAGAGGCGCATGGCGCGCGGGCGCTGTCTGGCGCTGTCTGGCG 80  
Db 1 GGCCCGGAGAGAGGCGCATGGCGCGCGGGCGCTGTCTGGCGCTGTCTGGCG 60  
Qy 81 TGGGCTGGACTCAGGAAGCGGAGTCCGAGGAGCGGCGGCTTATCAGGACCATGCG 140  
Db 61 TCGGCTGGACTCAGGAAGCGGAGTCCGAGGAGCGGCGGCTTATCAGGACCATGCG 120  
Qy 141 CCGAGGGTCACTACGTGCGCGCATCTGGTGGAGAGGACGCCGAGCTGGCGCTGGCG 200  
Db 121 CCGAGGGTCACTACGTGCGCGCATCTGGTGGAGAGGACGCCGAGCTGGCGCTGGCG 180  
Qy 201 GTGGCAGGGAGCGCTGCGCCTGTGGGATTCCACCATGTATCGGAGTGAGCTCTCAGCCA 260  
Db 181 GTGGCAGGGAGCGCTGCGCCTGTGGGATTCCACCATGTATCGGAGTGAGCTCTCAGCCA 240  
Qy 261 CCGTGGGCACTCAGCGGCGGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTC 320  
Db 241 CCGTGGGCACTCAGCGGCGGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTC 300  
Qy 321 CCGTGGGATGTCCAGTTTGGCCAGCTGACTTCCATGTCATCTTCTGGAGCTCGAGCG 380  
Db 301 CCGTGGGATGTCCAGTTTGGCCAGCTGACTTCCATGTCATCTTCTGGAGCTCGAGCG 360  
Qy 381 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 440  
Db 361 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 420  
Qy 441 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGACCTGTCTACCTACACTAAACATCA 500  
Db 421 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGACCTGTCTACCTACACTAAACATCA 480  
Qy 501 GCCCATCTCTCCAGGCGCTCCACATTTGAGTTTGAAGACGAGAGCTGTGGGTGAC 560  
Db 481 GCCCATCTCTCCAGGCGCTCCACATTTGAGTTTGAAGACGAGAGCTGTGGGTGAC 540  
Qy 561 TGGCTGGGGGTACATCAAGAGAGTGAAGCACTGCGCATCTCCCAACACCCCTCAGGAAGT 620  
Db 541 TGGCTGGGGGTACATCAAGAGAGTGAAGCACTGCGCATCTCCCAACACCCCTCAGGAAGT 600  
Qy 621 TCAGTGGCCATCATAAACAACTATATGCAACCACTCTTCTTCAAGTACAGTTTCG 680  
Db 601 TCAGTGGCCATCATAAACAACTATATGCAACCACTCTTCTTCAAGTACAGTTTCG 660  
Qy 681 CAAGGACATCTTTGGAGACATGTTTGTCTGGCAATGCCCAAGCGGAGGATGCCCTG 740  
Db 661 CAAGGACATCTTTGGAGACATGTTTGTCTGGCAATGCCCAAGCGGAGGATGCCCTG 720





Qy	501	GCCCATCTGTCTCCAGCCCTCCACATTTGAGTTTGAACCCGACACAGACTGCTGGGTGAC	560
Db	481	CCCCATCTGTCTCCAGCCCTCCACATTTGAGTTTGAACCCGACACAGACTGCTGGGTGAC	540
Qy	561	TGGCTGGGGGTACATCAAGAGAGATGAGGCACCTGCCATCTCCCCACACACCTCCAGGAAGT	620
Db	541	TGGCTGGGGGTACATCAAGAGAGATGAGGCACCTGCCATCTCCCCACACACCTCCAGGAAGT	600
Qy	621	TCAGGTGCCCATATAAACAACTCTATGTGCAACACACTTCTCTCAAGTACAGTTTCGG	680
Db	601	TCAGGTGCCCATATAAACAACTCTATGTGCAACACACTTCTCTCAAGTACAGTTTCGG	660
Qy	681	CAAGGACATCTTTGGAGACATGTTTGTCTGGCAATGCCAAGCGGGGAAAGATGCCCTG	740
Db	661	CAAGGACATCTTTGGAGACATGTTTGTCTGGCAAGCCCAAGCGGGGAAAGATGCCCTG	720
Qy	741	CTTTCGTGACTCAGGTGGACCCCTTGCGCTGTAAACAAGATGGACTGTGTATCAGATTGG	800
Db	721	CTTTCGTGACTCAGGTGGACCCCTTGCGCTGTAAACAAGATGGACTGTGTATCAGATTGG	780
Qy	801	AGTCGTGAGCTGGGGAGTGGCGTGTGTCGGGCCCAATCGGCCCGGTGTCTACACCAAT	860
Db	781	AGTCGTGAGCTGGGGAGTGGCGTGTGTCGGGCCCAATCGGCCCGGTGTCTACACCAAT	840
Qy	861	CAGCCACCACTTTGAGTGGATCCAGAAGCTGATGGCCCAAGAGTGGCATGTCCACGCCAGA	920
Db	841	CAGCCACCACTTTGAGTGGATCCAGAAGCTGATGGCCCAAGAGTGGCATGTCCACGCCAGA	900
Qy	921	CCCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGGCCGGT	980
Db	901	CCCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGGCCGGT	960
Qy	981	CTGAGCCTACCTTGAGCCCATCGAGCCTGGGGCCACTGCCAAGTCAGGCGCTTGGTCTCT	1040
Db	961	CTGAGCCTACCTTGAGCCCATCGAGCCTGGGGCCACTGCCAAGTCAGGCGCTTGGTCTCT	1020
Qy	1041	CTGTCTCTGTTGGTAAATAACACATTCAGTTGATGCT	1079
Db	1021	CTGTCTCTGTTGGTAAATAACACATTCAGTTGATGCT	1059

## RESULTAT 13

RES-001 13

US-09-904-786-256

; Sequence 256, Application US/09904786

; Publication No. US20030039969A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Garber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

Db 841 CAGCCACCACCTTGTAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGTCCAGCCAGA 900  
Qy 921 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGCCGGT 980  
Db 901 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGCCGGT 960  
Qy 981 CTGAGCTTACCTGAGCCCATGAGCCTGGGCCACTGCCAAGTCAAGGCCCTGTTCTCTT 1040  
Db 961 CTGAGCTTACCTGAGCCCATGAGCCTGGGCCACTGCCAAGTCAAGGCCCTGTTCTCTT 1020  
Qy 1041 CTGCTCTTCTTGGTAATAAACACATTCAGTTGATGCCT 1079  
Db 1021 CTGCTCTTCTTGGTAATAAACACATTCAGTTGATGCCT 1059

RESULT 14

US-09-906-646-256  
; Sequence 256, Application US/09906646  
; Publication No. US20030039971A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavain, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/906,646  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 256  
; LENGTH: 1100  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; US-09-906-646-256  
  
Query Match 97.7%; Score 1055.8; DB 9; Length 1100;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 21 GGGCCGGGAGAGAGGCCATGGCGCGCGGGCGCTGCTGTGCGGCTCTCTGCTGCG 80  
Db 1 GGGCCGGGAGAGAGGCCATGGCGCGCGGGCGCTGCTGTGCGGCTCTCTGCTGCG 60  
Qy 81 TCGGCTGGACTCAGGAAGCCGGAGTCCAGGAGGCGGCGCCCTTATCAGGACCATGG 140  
Db 61 TCGGCTGGACTCAGGAAGCCGGAGTCCAGGAGGCGGCGCCCTTATCAGGACCATGG 120  
Qy 141 CCGACGGGTTCATCAGTCGCGCATCTGGGTGGAGAGACGCCGAACCTCGGCGCTGGCC 200  
Db 121 CCGACGGGTTCATCAGTCGCGCATCTGGGTGGAGAGACGCCGAACCTCGGCGCTGGCC 180  
Qy 201 GTGGCAGGGAGCCCTGCGCCTGTGGGATTCGCCATGTATGCGGAGTGGAGCTCTCAGCCA 260  
Db 181 GTGGCAGGGAGCCCTGCGCCTGTGGGATTCGCCATGTATGCGGAGTGGAGCTCTCAGCCA 240  
Qy 261 CCGCTGGGCACTCAGCGGCGGCGCACTGCTTTGAACCTATAGTGACCTTAGTGATCCCTC 320  
Db 241 CCGCTGGGCACTCAGCGGCGGCGCACTGCTTTGAACCTATAGTGACCTTAGTGATCCCTC 300  
Qy 321 CCGGTGGATGGTCCAGTTTGGCCAGCTGACCTTCCATGCCATCCTTCTGGAGCCTCAGCG 380  
Db 301 CCGGTGGATGGTCCAGTTTGGCCAGCTGACCTTCCATGCCATCCTTCTGGAGCCTCAGCG 360  
Qy 381 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 440  
Db 361 CTACTACACCCGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGAATTC 420  
Qy 441 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCA 500  
Db 421 ACCCTATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCA 480  
Qy 501 GCCCATCTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC 560  
Db 481 GCCCATCTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGAC 540  
Qy 561 TGGCTGGGGGTACATCAAGAGGATGAGGCATGCGCATCTCCCCACACCCCTCCAGGAAGT 620  
Db 541 TGGCTGGGGGTACATCAAGAGGATGAGGCATGCGCATCTCCCCACACCCCTCCAGGAAGT 600  
Qy 621 TCAGGTCCCATCATAAACAACCTCTATGTGAACACCTCTTCTCAAGTACAGTTTCG 680  
Db 601 TCAGGTCCCATCATAAACAACCTCTATGTGAACACCTCTTCTCAAGTACAGTTTCG 660  
Qy 681 CAAGGACATCTTTGGAGACATGTTGTGCTGGCAATGCCAAGCGGAGGATGCCTG 740  
Db 661 CAAGGACATCTTTGGAGACATGTTGTGCTGGCAATGCCAAGCGGAGGATGCCTG 720  
Qy 741 CTTGGTGACTCAGGTGGACCCCTTGGCCCTGTAAACAAGATGGACTGTGTATCAGATTGG 800  
Db 721 CTTGGTGACTCAGGTGGACCCCTTGGCCCTGTAAACAAGATGGACTGTGTATCAGATTGG

Db	721	CTTCGGTGACTCAGGTGGACCCCTTGSCCTGTAAACAAGAAATGGACTGTGGTATCATGATGG	780
QY	801	AGTCGTGAGCTGGGGAGTGGGTGTGGTCGGCCCAATCGGCCCGTCTACACCAATAT	860
Db	781	AGTCGTGAGCTGGGGAGTGGGTGTGGTCGGCCCAATCGGCCCGTCTACACCAATAT	84
QY	861	CAGCCACCACCTTGAGTGGATCCAGAGCTGATGCCAGAGTGGCATGTCCCAGCCAGA	920
Db	841	CAGCCACCACCTTGAGTGGATCCAGAGCTGATGCCAGAGTGGCATGTCCCAGCCAGA	900
QY	921	CCCTCCTGGCCACTACTCTTTTCCCCTTCTCTGGGCTCTCCCACTCTGGGGCCGGT	980
Db	901	CCCTCCTGGCCACTACTCTTTTCCCCTTCTCTGGGCTCTCCCACTCTGGGGCCGGT	960
QY	981	CTGAGCCTACCTTGAGCCCATGCAAGCTGGGGCACTGCCAAGTCAGGCCCTGGTTCCTT	1040
Db	961	CTGAGCCTACCTTGAGCCCATGCAAGCTGGGGCACTGCCAAGTCAGGCCCTGGTTCCTT	1020
QY	1041	CTGTCTGTGTTGGTAAATAACACATATCCAGTTGATGCT	1079
Db	1021	CTGTCTGTGTTGGTAAATAACACATATCCAGTTGATGCT	1059

**RESULT 15**

US-09-906-700-256  
Sequence 256, Application US/0906700  
Publication No. US20030039972A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/906,700  
CURRENT FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547

```

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/230899
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 256
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-06/700-256

```

Query Match	97.7%;	Score 1055.8;	DB 9;	Length 1100;			
Best Local Similarity	99.8%;	Pred. No. 0;					
Matches 1057; Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	21	GGCCGCGGAGAGAGGCCATGGGCGCGCGCGGGCGCTGTCTGTGCGCGCTGCTGCTGC	80				
Db	1	GGCCGCGGAGAGAGGCCATGGGCGCGCGCGGGCGCTGTCTGTGCGCGCTGCTGCTGC	60				
QY	81	TGCGGCTGGACTCAGGAAGCCGAGTTCGACGAGAGCGCGCCCTTATCAGACCATTGCGG	140				
Db	61	TCGGCTGGACTCAGAAAGCCGGAGTTCGACGAGAGCGCGCGTTATCAGAACCATGCGG	120				
QY	141	CCGAGGGTCATCACGTCCGCATCGTGGGTGGAGAGGACGCCAATTCGGGGCTTGCGC	200				
Db	121	CCGAGGGTCATCACGTCCGCATCGTGGGTGGAGAGGACGCCAATTCGGGGCTTGCGC	180				
QY	201	GTGGCAGGGGAGCCTGCGCCCTGTGGGATTCACAGTATGCGGAGTGAGCCTGCTCAGCCA	260				
Db	181	GTGGCAGGGGAGCCTGCGCCCTGTGGGATTCACAGTATGCGGAGTGAGCCTGCTCAGCCA	240				
QY	261	CGCGTGGGCACCTACGCGCGCGCACTGCTTTGAAACCTATAGTAGCCTTAGTATCCCCTC	320				
Db	241	CGCGTGGGCACCTACGCGCGCGCACTGCTTTGAAACCTATAGTAGCCTTAGTATCCCCTC	300				
QY	321	CGGTTGGATGTCAGTTTGGCCAGCTTCCATGCCATCCTTCTGGAGCCTGCAGGC	380				
Db	301	CGGTTGGATGTCAGTTTGGCCAGCTTCCATGCCATCCTTCTGGAGCCTGCAGGC	360				
QY	381	CTACTACACCGTTACTTCGTATCGAATATCTATGAGCCCTCGGTACTGGGGAATTC	440				
Db	361	CTACTACACCGTTACTTCGTATCGAATATCTATGAGCCCTCGGTACTGGGGAATTC	420				
QY	441	ACCCATGACATTGCCCTTGGTGAAGCTGCTGCACCTGTCACTACACTAAACACATCCA	500				
Db	421	ACCCATGACATTGCCCTTGGTGAAGCTGCTGCACCTGTCACTACACTAAACACATCCA	480				
QY	501	GCCCATCTGTCTCAGGCCCTCCACATTTGAGTTTGAGAACCAGACAGACTGCTGGGTGAC	560				
Db	481	GCCCATCTGTCTCAGGCCCTCCACATTTGAGTTTGAGAACCAGACAGACTGCTGGGTGAC	540				
QY	561	TGGCTGGGGTACATCAAGAGGATGAGGCATTCGCCATCTCCCCACACCTCCAGGAAGT	620				
Db	541	TGGCTGGGGTACATCAAGAGGATGAGGCATTCGCCATCTCCCCACACCTCCAGGAAGT	600				
QY	621	TCAGGTGCCCATCAAACAACCTCTATGTGCAACACCTCTTCTCAAGTACAGTTTCG	680				
Db	601	TCAGGTGCCCATCAAACAACCTCTATGTGCAACACCTCTTCTCAAGTACAGTTTCG	660				



Qy	681	CAAGGACATCTTTGGAGACATGGTTTGTGCTGGCAATGCCAAGGCGGGAAGGATGCCTG	740
Db	661	CAAGGACATCTTTGGAGACATGGTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTG	720
Qy	741	CTTCGGTGACTCAGGTGGACCTTGGCCCTGTAAACAAGAAATGGACTGTGGTATCAGATTGG	800
Db	721	CTTCGGTGACTCAGGTGGACCTTGGCCCTGTAAACAAGAAATGGACTGTGGTATCAGATTGG	780
Qy	801	AGTCGTGAGCTGGGAGTGGGCTGTGCTGGCCCAATCGGCCCGGTGTCTACACCAATAT	860
Db	781	AGTCGTGAGCTGGGAGTGGGCTGTGCTGGCCCAATCGGCCCGGTGTCTACACCAATAT	840
Qy	861	CAGCCACCACCTTTGAGTGGATCCAGAAGCTGATGGCCCAAGAGTGGCATGTCCACGCCAGA	920
Db	841	CAGCCACCACCTTTGAGTGGATCCAGAAGCTGATGGCCCAAGAGTGGCATGTCCACGCCAGA	900
Qy	921	CCCTTCCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCCGGGCCGGT	980
Db	901	CCCTTCCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCCGGGCCGGT	960
Qy	981	CTGAGCCTACCTGAGCCCATGAGCCTGGGCCACTGCCAAGTCAAGGCCCTGGTTCTCTT	1040
Db	961	CTGAGCCTACCTGAGCCCATGAGCCTGGGCCACTGCCAAGTCAAGGCCCTGGTTCTCTT	1020
Qy	1041	CTGCTCTTCTTTGGTAATAAACACATTCAGTTGATGCCT	1079
Db	1021	CTGCTCTTCTTTGGTAATAAACACATTCAGTTGATGCCT	1059

Search completed: May 10, 2003, 05:00:59  
Job time : 219 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 03:18:27 ; Search time 2077 seconds

(without alignments)  
8429.139 Million cell updates/sec

Title: 05-09-787-844-1

Perfect score: 1081

Sequence: 1 aggagggcagggggcgctca.....acattccagttgacgtcgc 1081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_estl.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: gb\_gss.\*

18: em\_gss\_hum.\*

19: em\_gss\_inv.\*

20: em\_gss\_pln.\*

21: em\_gss\_vrt.\*

22: em\_gss\_fun.\*

23: em\_gss\_mam.\*

24: em\_gss\_mus.\*

25: em\_gss\_other.\*

26: em\_gss\_pro.\*

27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	906.8	83.9	936	9 AL578261	AL578261 AL578261
2	862.4	79.8	876	9 AL555870	AL555870 AL555870
3	777.2	71.9	871	13 BI334059	BI334059 602997320
4	760.4	70.3	777	13 BI259237	BI259237 602970095
5	733.2	67.8	878	13 BI827608	BI827608 603073575
c 6	624.8	57.8	633	9 AI126185	AI126185 qd81h05.x

7	567.6	52.5	705	13	BM019026
8	551.4	51.0	682	13	BM017444
9	549.8	50.9	1219	13	BM553333
10	523.4	48.4	824	13	BM018961
c 11	523	48.4	591	9	AA161187
12	517.4	47.9	1286	14	BM911057
13	501.6	46.4	825	13	BI257649
14	486.2	45.0	1050	11	AK006271
15	483.6	44.7	624	13	BM018436
16	478	44.2	625	12	EG717268
17	474	43.8	628	12	EG717347
18	468	43.3	479	9	AA448095
19	452	41.8	484	14	BQ252128
c 20	451.4	41.8	514	9	AA620757
21	447.4	41.4	510	14	BM759973
c 22	441.6	40.9	462	9	AI016399
c 23	412.2	38.1	487	9	AA525010
c 24	383	35.4	399	9	AI278269
c 25	383	35.4	412	9	AA448166
c 26	381.4	35.3	402	9	AA729680
c 27	379.8	35.1	406	9	AI150137
c 28	361.4	33.4	383	9	AA447596
c 29	351.2	32.5	454	13	BI257498
30	337	31.2	1089	11	AK016509
c 31	326.4	30.2	411	9	AA149942
c 32	318.2	29.4	373	9	AA401115
c 33	311.2	28.8	331	9	AI217989
c 34	306.4	28.3	334	9	AA552523
c 35	306.2	28.3	475	10	AW971363
c 36	295.2	27.3	542	9	AI326140
c 37	288.4	26.7	351	9	AA860356
c 38	269	24.9	950	10	BE546936
c 39	236.2	21.9	359	9	AI926900
c 40	223.6	20.7	582	13	BG987202
c 41	222	20.5	482	12	BF405556
c 42	222	20.5	624	12	BF404444
c 43	222	20.5	625	12	BF403421
c 44	222	20.5	690	12	BF405109
c 45	212	19.6	431	12	BF405483

#### ALIGNMENTS

RESULT 1  
LOCUS AL578261/c 936 bp mRNA linear EST 16-FEB-2001  
DEFINITION AL578261 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CSODK002YM07 3  
prime, mRNA sequence.  
ACCESSION AL578261  
VERSION AL578261.1 GI:12942167  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Li.W.B., Gruber.C., Jessee.J. and Polayes.D.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 936)  
AUTHORS Full-length cDNA libraries and normalization  
TITLE Unpublished (2001)  
JOURNAL Contact: Genoscope  
COMMENT Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. .936  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODK002YM07"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [filang@lifetech.com](mailto:filang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

```

BASE COUNT
ORIGIN
206 a 256 c 288 g 180 t 6 others
http://fulllength.invitrogen.com"

```

Query Match	83.9%	Score 906.8	DB 9	Length 936
Best Local Similarity	98.5%	Pred. No. 1.3e-223		
Matches 909	Conservative 5	Mismatches 9	Indels 0	Gaps 0
QY 107	CGCAGGAGCGCGCCCTTATCAGGACCATGCGCGCGACGGGTCTATCAGTCCGCATCG	166		
Db				
QY 936	CGCAGGAGCGCGCGCTTATCAGGACCATGCGCGCGACGGGTCTATCAGTCCGCATCG	877		
Db				
QY 167	TGGGTGGAGAGGACGCGGACCTCGGCGTGGCGGTGGCAGGGAGCCTGGCCTGTGGG	226		
Db				
QY 876	TGGGTGGAGAGGACGCGGACCTCGGCGTGGCGGTGGCAGGGAGCCTGGCCTGTGGG	817		
Db				
QY 227	ATTTCCACGTATGCGGAGTGAGCTGCTCAGGCACCGCTGGGCATCTCAGGCGGGCAGCT	286		
Db				
QY 816	ATTTCCACGTATGCGGAGTGAGCTGCTCAGGCACCGCTGGGCATCTCAGGCGGGCAGCT	757		
Db				
QY 287	GCCTTTGAACCTATAGTGACCTTACTGTATCCCTCCGGGTGGATGCTCCAGTTTGGCCAGC	346		
Db				
QY 756	GCCTTTGAACCTATAGTGACCTTACTGTATCCCTCCGGGTGGATGCTCCAGTTTGGCCAGC	697		
Db				
QY 347	TGACTTCCATGTCATCCCTCTCGGAGCTGAGGCTACTACACCGTTACTTCTGATATCGA	406		
Db				
QY 696	TRACTTCCATGTCATCCCTCTCGGAGCTGAGGCTACTACACCGTTTWTTCGATATCGA	637		
Db				
QY 407	ATATCTATCTGAGCCCTCGCTACCTTGGGGAATTCACCCATATGACATTCGCTTGGTGAAGC	466		
Db				
QY 636	ATATCTATCTGAGCCCTCGCTACCTTGGGGAATTCACCCATATGACATTCGCTTGGTGAAGC	577		
Db				
QY 467	TGCTTGCACCTGTCTACCTTACACTAAACATCCAGCCCATCTGTCTCCAGGCTCCACAT	526		
Db				
QY 576	TGCTTGCACCTGTCTACCTTACACTAAACATCCAGCCCATCTGTCTCCAGGCTCCACAT	517		
Db				
QY 527	TTGAGCTTTGAGAACGGGACAGACTGCTGGGTGACTGCTGGGGTACATCAAGAGGATG	586		
Db				
QY 516	TTGAGCTTTGAGAACGGGACAACTGCTGGGTAACTGCTGGGGTACATCAAGAGGATG	457		
Db				
QY 587	AGGCACTGCCATCTCCACACACCTCCAGGAGTTTCAAGTTCAGTTCGCCATCAAACTCTA	646		
Db				
QY 456	AGGCACTGCCATCTCCACACACCTCCAGGAGTTTCAAGTTCGCCATCAAACTCTA	397		
Db				
QY 647	TGTGAACCCCTCTTCTCAAGTACAGTTTCCGAGGACATCTTTGGAGACATGGTTT	706		
Db				
QY 396	TGTGAACCCCTCTTCTCAAGTACAGTTTCCGAGGACATCTTTGGAGACATGGTTT	337		
Db				
QY 707	GTGCTGGGATGCCAAGGCGGGAAGATGCCCTGCTTCGTGACTCAGGTGGACCTTGG	766		
Db				
QY 336	GTGCTGGCAATGCCAAGGCGGGAAGGATGCCCTGCTTCGGTAACTCAGGTGGACCTTGG	277		
Db				
QY 767	CCTGTACAGAATGGACTGTGTATCAGATTGGAGTCTGTGAGCTGGGAGTGGCTGTG	826		
Db				
QY 276	CCTGTACAGAATGGACTGTGTATCAGATTGGAGTCTGTGAGCTGGGAGTGGCTGTG	217		
Db				
QY 827	GTGCGCCCAATCGCGCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGA	886		
Db				
QY 216	GTGCGCCCAATCGCGCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGA	157		
Db				
QY 887	AGCTGATGGCCCCAGATGGCATGTCCAGCCAGACCCCTCTCGGCACTACTCTTTTTC	946		
Db				
QY 156	AGCTGATGGCCCCAGATGGCATGTCCAGCCAGACCCCTCTCGGCACTACTCTTTTTC	97		
Db				
QY 947	CTCTTCTCTGGGCTCTCCCACTCTCTGGGCGGGTCTGAGCCCTACTGAGCCCATCGAGCC	1006		
Db				

Db	96	CTCTTCHCTGGGCTCTCCCMCTCCTGGGGCCGGTCTTGAGCCCTACCTGAGCCCATGAGCC 37
QY	1007	TGGGGCCACTGCCAAGCTAGGCC 1029 
Db	36	TGGGGCCACTGCCAAGCAGGCC 14 
RESULT 2		
AL555870		
LOCUS		
DEFINITION		876 bp mRNA linear EST 16-FEB-2001
ACCESSION	AL555870	LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DK002YM07 5
VERSION	AL555870	prime, mRNA sequence.
KEYWORDS	AL555870.1	GI:12898012
SOURCE	EST.	
ORGANISM	human.	
REFERENCE	Homo sapiens	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1. (bases 1 to 876)	
JOURNAL	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.	
COMMENT	Full-length cDNA libraries and normalization unpublished (2001)	
FEATURES	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.	
SOURCE	Location/Qualifiers 1..876 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0DK002YM07" /clone_lib="LTI_NFL006_PL2" /tissue_type="placenta" /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"	
BASE COUNT	171 a	251 c 266 g 188 t
ORIGIN		

Query Match	79.8%	Score 862.4	DB 9	Length 876
Best Local Similarity	99.8%	Pred. No. 3.8e-212		
Matches 874	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	34	GAGGCCATGGCGCGCGCGGGCGCGTGTCTGTGGCGCTGCTCTGCTGCTCGGCGCTGGACTC	93	
Db				
	1	GAGGCCATGGCGCGCGCGGGCGCGTGTCTGTGGCGCTGCTCTGCTGCTCGGCGCTGGACTC	60	
QY	94	AGGAAGCCGGAGTGCAGGAGCGCGGCCCT-TATCAGGACCATCGGCCGACGGGTCAT	152	
Db				
	61	AGGAAGCCGGAGTGCAGGAGCGCGCGCTGTATCAGGACCATCGGCCGACGGGTCAT	120	
QY	153	CACGTCGCGCATCTGGTGGAGAGGACGCCGAACCTCGGGCGTTGGCGTGCGCAGGGGAG	212	
Db				
	121	CACGTCGCGCATCTGGGTTGGAGAGGACGCCGAACCTCGGGCGTTGGCGTGCGCAGGGGAG	180	
QY	213	CCTCGCGCTCTGGGATTCGCCACGTTATCGGAGTGAGCGCTGCTCAGCCACCGCTGGGCAC	272	
Db				
	181	CCTCGCGCTCTGGGATTCGCCACGTTATCGGAGTGAGCGCTGCTCAGCCACCGCTGGGCAC	240	
QY	273	CAGGGCGGCGCATGCTTTTGAACCTATAGTAGCTTGTAGTCCCTCCCGGGTGGATGGT	332	
Db				
	241	CAGGGCGGCGCATGCTTTTGAACCTATAGTAGCTTGTAGTCCCTCCCGGGTGGATGGT	300	
QY	333	CCAGTTTGGCCAGCTGACTTCCATGGCATCCCTTCTGAGGCGCTCGAGGCGCTACTACACCG	392	

```
Db 301 CCAGTTTGGCAGCTGACTTCCATCCCATCTTCTGGAGCCTGCGAGGCTACTACACCG 360
Qy TTACTTGTATCGAATATCTATCTAGCCCTCGCTACTCTGGGAAATTCACCCCTATGACAT 452
Db 361 TTACTTGTATCGAATATCTATCTAGCCCTCGCTACTCTGGGAAATTCACCCCTATGACAT 420
Qy 453 TGCCTTGGTGAAGCTGTCTGACCTGTCACTACATCACTAAACACATCCAGCCCATCTGTCT 512
Db 421 TGCCTTGGTGAAGCTGTCTGACCTGTCACTACATCACTAAACACATCCAGCCCATCTGTCT 480
Qy 513 CCAGGCTCCACATTTAGTTTGAAGCCGACAGACTGCTGGTGACTGCTGGGGTA 572
Db 481 CCAGGCTCCACATTTAGTTTGAAGCCGACAGACTGCTGGTGACTGCTGGGGTA 540
Qy 573 CATCAAGAGATGAGGACTGCTGCTCCATCTCCACACCTCCAGGAAGTTCAAGTCCGCAT 632
Db 541 CATCAAGAGATGAGGACTGCTGCTCCATCTCCACACCTCCAGGAAGTTCAAGTCCGCAT 600
Qy 633 CATAAACAACCTCTATGTGCAACACCTCTCTCTCAAGTACAGTTTCCGCAAGGACATCTT 692
Db 601 CATAAACAACCTCTATGTGCAACACCTCTCTCTCAAGTACAGTTTCCGCAAGGACATCTT 660
Qy 693 TGGAGACATGTTTGTCTGCTGCAATGCCATCTCCACACCTCCAGGAAGTTCAAGTCCGCAT 752
Db 661 TGGAGACATGTTTGTCTGCTGCAATGCCATCTCCAGGAAGTTCAAGTCCGCAT 720
Qy 753 AGGTGGACCTTGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCGTGAGCTG 812
Db 721 AGGTGGACCTTGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCGTGAGCTG 780
Qy 813 GGGAGTGGGCTGTGGTGGGCCAATCGGCCGGTGTCTTACACCAATATACGCCACCACTT 872
Db 781 GGGAGTGGGCTGTGGTGGGCCAATCGGCCGGTGTCTTACACCAATATACGCCACCACTT 840
Qy 873 TGACTGGTCCAGAGCTGATGGCCGACAGTGGCAT 908
Db 841 TGACTGGTCCAGAGCTGATGGCCGACAGTGGCAT 876
```

## RESULT 3

```
BI334059 871 bp mRNA linear .EST 30-JUL-2001
LOCUS 602997320F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139509 5',
DEFINITION mRNA sequence.
ACCESSION BI334059
VERSION BI334059.1 GI:15018703
KEYWORDS EST.
SOURCE human.
```

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 871)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

## AUTHORS

Contact: Robert Strausberg, Ph.D.

## JOURNAL

Email: cgapbs-r@mail.nih.gov

## COMMENT

Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LUAM1343 row: h column: 06  
High quality sequence stop: 826.

## FEATURES

Location/Qualifiers  
1..871  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:5139509"  
/clone\_lib="NIH\_MGC\_12"  
/tissue\_type="cervical carcinoma cell line"

```
BASE COUNT 185 a 258 c 224 g 204 t
ORIGIN
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
```

```
Query Match 71.9%; Score 777.2; DB 13; Length 871;
Best Local Similarity 97.9%; Pred. No. 4e-190;
Matches 842; Conservative 0; Mismatches 8; Indels 10; Gaps 5;
Qy 209 GGAGCCTCGCCTGTGGGATTCACACGATCGGAGTGTGAGCTGTCTGAGCCCTGAGCCCTGGG 268
Db 1 GGAGCCTCGCCTGTGGGATTCACACGATCGGAGTGTGAGCTGTCTGAGCCCTGAGCCCTGGG 60
Qy 269 CACTCACGGCGGCGACTGTCTTGAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGA 328
Db 61 CACTCACGGCGGCGACTGTCTTGAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGA 114
Qy 329 TGGTCCAGTTTGGCCAGCTGACTTCCATGCTGCTCTGAGCCCTGAGCCCTGAGCCCTACTACA 388
Db 115 TGGTCCAGTTTGGCCAGCTGACTTCCATGCTGCTCTGAGCCCTGAGCCCTGAGCCCTACTACA 174
Qy 389 CCGCTTACTTCGTATCGAATATCTATCTGAGCCCTGCTGCTGAGCCCTGAGCCCTGAGCCCTATG 448
Db 175 CCGCTTACTTCGTATCGAATATCTATCTGAGCCCTGCTGCTGAGCCCTGAGCCCTATG 234
Qy 449 ACATTGCTTGGTGAAGCTGTCTGCACTGTGCACTTACACTAAACACATCCAGCCCTAT 508
Db 235 ACATTGCTTGGTGAAGCTGTCTGCACTGTGCACTTACACTAAACACATCCAGCCCTAT 294
Qy 509 GTCTCCAGCCCTCCACATTTGAGTTTGAACCGGAGACTGTCTGAGCCCTGAGCCCTGAGCCCTGGG 568
Db 295 GTCTCCAGCCCTCCACATTTGAGTTTGAACCGGAGACTGTCTGAGCCCTGAGCCCTGAGCCCTGGG 354
Qy 569 GGTACATCAAGAGAGATGAGGACTGCGCATCTCCACACCTCTCCAGGAAGTTTCAAGTCG 628
Db 355 GGTACATCAAGAGAGATGAGGACTGCGCATCTCCACACCTCTCCAGGAAGTTTCAAGTCG 414
Qy 629 CCATCATAAACAACCTCTATGTGCAACACCTCTTCTCTCAAGTACAGTTTCCGCAAGGACA 688
Db 415 CCATCATAAACAACCTCTATGTGCAACACCTCTTCTCTCAAGTACAGTTTCCGCAAGGACA 474
Qy 689 TCTTTGGAGACATGGTTTGTCTGGCAATGCCAAGCGGGAAGGATGCTCTGCTCGGTG 748
Db 475 TCTTTGGAGACATGGTTTGTCTGGCAATGCCAAGCGGGAAGGATGCTCTGCTCGGTG 534
Qy 749 ACTCAGTGGACCTTGGCCTGTAAACAAGATGGACTGTGGTATCAGATTGG-AGTCGTG 807
Db 535 ACTCAGTGGACCTTGGCCTGTAAACAAGATGGACTGTGGTATCAGATTGGCAGTCGTG 594
Qy 808 AGCTGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCAC 867
Db 595 AGCTGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCAC 654
Qy 868 CACTTTGAGTGGATCCAGAAGCTGATGCCAGAGTGGCATGTCCACGACAGACCCCTCC 927
Db 655 CACTTTGAGTGGATCCAGAAGCTGATGCCAGAGTGGCATGTCCACGACAGACCCCTCC 714
Qy 928 TGGCCACTACTCTTTTTCCTCTCTGCGCTCTCCACCTCTGCGGCGCGGCTGTAGGCC 987
Db 715 TGGCC-GTACTCTTTTTCCTCTCTGCGCTCTCCACCTCTGCGGCGCGGCTGTAGGCC 773
Qy 988 TACTGAGCCCATGAGCCTGGGCGCACTGCCAAGTCAAGCCCTGGTGTCTCTTCTGTCT 1047
Db 774 TAACTGAGCCCATGAGCCTGGGCGCACTG-CAAGTCCAGGCGCTGGTCTCTCTTCTGACTT 832
Qy 1048 GTTTGGTAAATAACACATTC 1067
Db 833 G-TAGGTAATAAAAAACATTC 851
```

```

RESULT 4
BI259237 777 bp mRNA linear EST 17-JUL-2001
LOCUS 602970095F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5109695 5',
DEFINITION mRNA sequence.
ACCESSION BI259237
VERSION BI259237.1 GI:14816376
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL 1 (bases 1 to 777)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1265 row: m column: 24
High quality sequence stop: 777.
Location/Qualifiers
1. 777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5109695"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT 159 a 224 c 224 g 170 t
ORIGIN
Query Match 70.3%; Score 760.4; DB 13; Length 777;
Best Local Similarity 99.7%; Pred. No. 8.4e-186;
Matches 772; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 91 CTCAGAGCCGAGTCGAGGAGGCGCGCCCTTATCAGGACATGCGCGACGGGTC 150
Db 1 CTCAGAGCCGAGTCGAGGAGGCGCGCCCTTATCAGGACATGCGCGACGGGTC 60
QY 151 ATCAGCTCGCGCATCGTGGGTGGAGAGGACGCCGAACCTCGGCGTGGCGGAGGG 210
Db 61 ATCAGCTCGCGCATCGTGGGTGGAGAGGACGCCGAACCTCGGCGTGGCGGAGGG 120
QY 211 AGCCTGCGCCTGGGATGCCACGTATGCGGAGTGAGCGTCTCAGCCACCGCTGGGCA 270
Db 121 AGCCTGCGCCTGGGATGCCACGTATGCGGAGTGAGCGTCTCAGCCACCGCTGGGCA 180
QY 271 CTCAGCGGGCGGCACCTGCTTGAACACCTATAGTAGACCTTAGTGATCCCTCGCGGTGGATG 330
Db 181 CTCAGCGGGCGGCACCTGCTTGAACACCTATAGTAGACCTTAGTGATCCCTCGCGGTGGATG 240
QY 331 GTCCAGTTTGGCCAGCTCAGTCCATGCTCCATGCTCTGAGCGCTGACGCGCTACTACACC 390
Db 241 GTCCAGTTTGGCCAGCTCAGTCCATGCTCCATGCTCTGAGCGCTGACGCGCTACTACACC 300
QY 391 CGTTACTTCGTATGGAATATCTATCTAGCGCTCGCTACCTCGGGAATTCACCTATGAC 450
Db 301 CGTTACTTCGTATGGAATATCTATCTAGCGCTCGCTACCTCGGGAATTCACCTATGAC 360
QY 451 ATTGCTTGGTGAAGCTGTGTCACCTGTACCTACACTAACACATCCAGCCCATCTGT 510
Db 361 ATTGCTTGGTGAAGCTGTGTCACCTGTACCTACACTAACACATCCAGCCCATCTGT 420

```

```

QY 511 CTCAGGCGCTCCACATTTGAGTTTGAGAACCGACAGACTGCTGGTGACTGGCTGGGG 570
Db 421 CTCAGGCGCTCCACATTTGAGTTTGAGAACCGACAGACTGCTGGTGACTGGCTGGGG 480
QY 571 TACATCAAGAGAGATGAGCACTGCCATCTCCCCACACACCTCCAGGAAGTTCAAGTCGCC 630
Db 481 TACATCAAGAGAGATGAGCACTGCCATCTCCCCACACACCTCCAGGAAGTTCAAGTCGCC 540
QY 631 ATCAATAACAACCTCTATGTGCAACACACCTCTTCTCAAGTACAGTTTCGCGAAGGACATC 690
Db 541 ATCAATAACAACCTCTATGTGCAACACACCTCTTCTCAAGTACAGTTTCGCGAAGGACATC 600
QY 691 TTTGAGACATGTTTGTGTGGCAATGCCAAGCGGGGAAGCATGCTCTTCGTTCCGTGAC 750
Db 601 TTTGAGACATGTTTGTGTGGCAATGCCAAGCGGGGAAGCATGCTCTTCGTTCCGTGAC 660
QY 751 TCAGGTGACCTTGGCCTGTAAACAAGATGAGTGTGGTATCAGATTGGAGTCGTGAGC 810
Db 661 TCAGGTGACCTTGGCCTGTAAACAAGATGAGTGTGGTATCAGATTGGAGTCGTGAGC 720
QY 811 TGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGGTGTCTACACCAATATCAGC 864
Db 721 TGGGAGTGGGCTGTGGTGGCG-CCATCGGCCCGGTGTCTACACCAATATCAGC 773

RESULT 5
BI27608 878 bp mRNA linear EST 04-OCT-2001
LOCUS 603073575F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5165562 5',
DEFINITION mRNA sequence.
ACCESSION BI27608
VERSION BI27608.1 GI:15939145
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM11411 row: e column: 19
High quality sequence stop: 828.
Location/Qualifiers
1. 878
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5165562"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT 172 a 250 c 256 g 200 t
ORIGIN

```

3' similar to SW:PSS8\_HUMAN Q16651 PROSTASIN PRECURSOR ; , mRNA sequence.

AII126185  
AII126185.1 GI:3594699  
EST.  
human.

Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 633)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabps.r@mail.nih.gov  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo,  
Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 912 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 455.  
Location/Qualifiers  
1..633  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:1735929"  
/glone\_lib="Soares\_testis\_NHT"  
/sex="male"  
/lab\_host="DH10B"  
/note=Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc., and primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGAGCGGCCCATTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization to Cot5, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
151 a 164 c 185 g 131 t 2 others

ch 57.8%; Score 624.8; DB 9; Length 633;  
l similarity 99.4%; Pred. No. 8.9e-151;  
626; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

ACATTGTCCTTGTGTGAGCTGTGCACCCTGCACCTACACTAAACACATCCGCCCATC 507  
ACATTGTCCTTGTGTGAGCTGTGCACCCTGCACCTAACACATCCAGCCCACATC 574  
ACATTGTCCTTGTGTGAGCTGTGTGAGAACCGGACAGACTGCTGGGTGACTGGCTGG 567  
GGTATCATTAAGAGGATGAGGCACCTGCCATCTCCCCACACCCCTCCAGGAAGTTCAGGTC 627  
GGTATCATTAAGAGGATGAGGCACCTGCCATCTCNCCACACCCCTCCAGGAAGTTCAGGTC 454  
CCCATATAACAACACTCTATGTGCAACCACTCTTCCTCAAGTACAGTTCGCCGAAGAC 687  
CCCATATAACAACACTCTATGTGCAACCACTCTTCCTCAAGTACAGTTCGCCGAAGAC 394  
CTCTTTGGAGACATGGTTTGTGTCGGCAATGCCAAGCGGGAAGGATGCTGCTCGGT 747  
CTCTTTGGAGACATGGTTTGTGTCGGCAATGCCAAGCGGGAAGGATGCTGCTCGGT 334  
ACTCAGGTGGACCCCTGGCCCTGTAAACAAGATGGACTGTGGTATCAGATGGAGTCGT 807

Db	333	GACTCAGGTGGACCCCTTGGCCCTGTAAACAGAAATGGACTGTGGTATCAGATTGGAGTCGTG	274
Qy	808	AGCTGGGGAGTGGGTGTGGTGGCCCAATCGGCCCGGTGTCTACACAAATATCAGCCAC	867
Db	273	AGCTGGGGAGTGGGTGTGGTGGCCCAATCGGCCCGGTGTCTACACAAATATCAGCCAC	214
Qy	868	CACTTTGAGTGGATCCAGAAAGCTGATGCCCCAGAGTGGCATGTCCAGCCAGACCCCTCC	927
Db	213	CACTTTGAGTGGATCCAGAAAGCTGATGCCCCAGAGTGGCATGTCCAGCCAGACCCCTCC	154
Qy	928	TGGCCACTACACTCTTTTCCCTTCTCTGGGCTCTCCCACTCTCTGGGGCCGGTCTGAGCC	987
Db	153	TGGCCCGTACACTCTTTTCCCTTCTCTGGGCTCTCCCACTCTCTGGGGCCGGTCTGAGCC	94
Qy	988	TACCTGAGCCCATCGACGCTGGGGCCACTGCCAAGTCTAGGCCCTGGTTCCTCTGTGCTT	1047
Db	93	TACCTGAGCCCATCGACGCTGGGGCCACTGCCAAGTCTAGGCCCTGGTTCCTCTGTGCTT	34
Qy	1048	GTTTGGTAATAAACACATATCCAGTTGATGC	1077
Db	33	GTTTGGTAATAAACACATATCCAGTTGATGC	4
RESULT	7		
LOCUS	BM019026		
DEFINITION	603647010F1 NIH_MGC_98 Homo sapiens cDNA linear	705 bp mRNA	EST 30-OCT-2001
ACCESSION	BM019026		mRNA sequence.
VERSION	BM019026.1	GI:16533380	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 705)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgabbs-re@mail.nih.gov">cgabbs-re@mail.nih.gov</a>		
	Tissue procurement: ATCC		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Plate: LLCM1896	row: c	column: 04
	High quality sequence start: 7		
	High quality sequence stop: 675.		

FEATURES	SOURCE
1. <b>High Accuracy:</b> The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.	1. <b>High Accuracy:</b> The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.
2. <b>Scalability:</b> The model is designed to scale efficiently, handling large volumes of data and complex tasks without significant performance degradation.	2. <b>Scalability:</b> The model is designed to scale efficiently, handling large volumes of data and complex tasks without significant performance degradation.
3. <b>Interpretability:</b> The model's decisions are transparent and explainable, allowing users to understand the underlying reasoning and identify potential biases or errors.	3. <b>Interpretability:</b> The model's decisions are transparent and explainable, allowing users to understand the underlying reasoning and identify potential biases or errors.
4. <b>Robustness:</b> The model is highly robust to noise and outliers, maintaining stable performance even in challenging or adversarial environments.	4. <b>Robustness:</b> The model is highly robust to noise and outliers, maintaining stable performance even in challenging or adversarial environments.
5. <b>Efficiency:</b> The model is optimized for fast execution, minimizing latency and maximizing throughput, making it suitable for real-time applications.	5. <b>Efficiency:</b> The model is optimized for fast execution, minimizing latency and maximizing throughput, making it suitable for real-time applications.
6. <b>Flexibility:</b> The model is adaptable to various input formats and tasks, allowing it to be easily integrated into different systems and workflows.	6. <b>Flexibility:</b> The model is adaptable to various input formats and tasks, allowing it to be easily integrated into different systems and workflows.
7. <b>Security:</b> The model is designed with robust security measures, ensuring data privacy and protecting against unauthorized access or manipulation.	7. <b>Security:</b> The model is designed with robust security measures, ensuring data privacy and protecting against unauthorized access or manipulation.
8. <b>Customizability:</b> The model offers a high degree of customization, allowing users to tailor the model's behavior and parameters to specific use cases and requirements.	8. <b>Customizability:</b> The model offers a high degree of customization, allowing users to tailor the model's behavior and parameters to specific use cases and requirements.
9. <b>Integration:</b> The model is seamlessly integrated with existing systems and infrastructure, facilitating easy adoption and deployment.	9. <b>Integration:</b> The model is seamlessly integrated with existing systems and infrastructure, facilitating easy adoption and deployment.
10. <b>Support:</b> The model is backed by a dedicated support team, providing users with timely assistance and guidance throughout the implementation and usage process.	10. <b>Support:</b> The model is backed by a dedicated support team, providing users with timely assistance and guidance throughout the implementation and usage process.

LOCUS BM017444 682 bp mRNA linear EST 30-OCT-2001  
DEFINITION G03644544F1 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5426338 5',  
mRNA sequence.  
ACCESSION BM017444  
VERSION BM017444.1 GI:16531798  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 682)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be



found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM1890 row: 9 column: 11  
 High quality sequence stop: 672.  
 Location/Qualifiers

## FEATURES

source  
 1. 682  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5426338"  
 /clone\_lib="NIH\_MGC\_98"  
 /tissue\_type="astrocytoma grade IV, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."  
 BASE COUNT 135 a 204 c 197 g 146 t

Query Match 51.0%; Score 551.4; DB 13; Length 682;  
 Best Local Similarity 97.9%; Pred. No. 8.5e-132;  
 Matches 569; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
 QY 120 GCCCTTATCAGGACCATCGCGCCGACGGTTCATCACGTCGCGCATCTGGGTGGAGAGGA 179  
 DB 102 GACATCCGAGGACCATCGCGCCGACGGTTCATCACGTCGCGCATCTGGGTGGAGAGGA 161  
 QY 180 CGCGAACTCGGGGTTGGCC-~~GTGGCAGGGAGCCCTCGCGCTCTGGGATTC~~CCACGAT 238  
 DB 162 CGCGAACTCGGGGTTGGCCGTGGCAGGGAGCCCTCGCGCTCTGGGATTCACGAT 221  
 QY 239 GCGAGTAGGCTCTCAGCCACCGCTGGGCACTACGCGGGCGCACTGCTTTGAAACCT 298  
 DB 222 GCGAGTAGGCTCTCAGCCACCGCTGGGCACTACGCGGGCGCACTGCTTTGAAACCT 281  
 QY 299 ATAGTGACCTTAGTGATCCCTCGGGTGGATGGTCCAGTTGGCCAGCTGACTTCCATGC 358  
 DB 282 ATAGTGACCTTAGTGATCCCTCGGGTGGATGGTCCAGTTGGCCAGCTGACTTCCATGC 341  
 QY 359 CATCTTCTGAGGCTCGAGGCTTACTACACCCGTTACTTCTGATCGAATATCTATCTGA 418  
 DB 342 CATCTTCTGAGGCTCGAGGCTTACTACACCCGTTACTTCTGATCGAATATCTATCTGA 401  
 QY 419 GCGCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTTGGTGAAGCTGCTGCACCTG 478  
 DB 402 GCGCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTTGGTGAAGCTGCTGCACCTG 461  
 QY 479 TCACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTTGA 538  
 DB 462 TCACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTTGA 521  
 QY 539 ACCGACAGACTGCTGGGTGACTGGCTGGGGTATACAAAGAGGATGAGCACTGCCAT 598  
 DB 522 ACCGACAGACTGCTGGGTGACTGGCTGGGGTATACAAAGAGGATGAGCACTGCCAT 581  
 QY 599 CTCCACACACCTCCAGGAGTTTCAGTTCGCCATCATACAACTCTATGTGCAACACC 658  
 DB 582 CTCCACACACCTCCAGGAGTTTCAGTTCGCCATCATACAACTCTATGTGCAACACC 641  
 QY 659 TCTTCTCAAGTACAGTTTCCGGAAGGACATCTTTGGAGAC 699  
 DB 642 TCTTCTCAAGTACAGTTTCCGGAAGGACATCTCGGGAGAC 682

## RESULT 9

BM553333

LOCUS

DEFINITION

BM553333 1219 bp mRNA linear EST 20-FEB-2002  
 AGENCOURT\_6558382 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5742958  
 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

## ORGANISM

BM553333  
 BM553333.1 GI:18791973  
 EST.  
 human.

## Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 1219)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM12761 row: 0 column: 23  
 High quality sequence stop: 553.  
 Location/Qualifiers

FEATURES  
source

1. 1219  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5742958"  
 /clone\_lib="NIH\_MGC\_119"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source normal medulla from  
 anonymous male age 27. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.3 kb, insert size range  
 0.9-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 013. Note:  
 this is a NIH\_MGC Library."  
 BASE COUNT 225 a 393 c 354 g 246 t 1 others

## ORIGIN

Query Match 50.9%; Score 549.8; DB 13; Length 1219;  
 Best Local Similarity 91.1%; Pred. No. 2.6e-131;  
 Matches 606; Conservative 0; Mismatches 57; Indels 2; Gaps 2;  
 QY 120 GCCCTTATCAGGACCATCGCGCCGACGGTTCATCACGTCGCGCATCTGGGTGGAGAGGA 179  
 DB 57 GACCATCCGAGGACCATCGCGCCGACGGTTCATCACGTCGCGCATCTGGGTGGAGAGGA 116  
 QY 180 CGCGAACTCGGGGTTGGCCGTGGCAGGGAGCTGCGCTCTGGGATTCACGATG 239  
 DB 117 CGCGAACTCGGGGTTGGCCGTGGCAGGGAGCTGCGCTCTGGGATTCACGATG 176  
 QY 240 CGGAGTAGCCTGCTCAGCCACCGCTGGGCACTCACGGCGCGCACTGCTTTGAAACCTA 299  
 DB 177 CGGAGTAGCCTGCTCAGCCACCGCTGGGCACTCACGGCGCGCACTGCTTTGAAACCTA 236  
 QY 300 TAGTGACCTTAGTGATCCCTCGGGTGGATGGTCCAGTTGGCCAGCTGACTTCCATGCC 359  
 DB 237 TAGTGACCTTAGTGATCCCTCGGGTGGATGGTCCAGTTGGCCAGCTGACTTCCATGCC 296  
 QY 360 ATCTTCTGGAGCTCGAGGCTTACTACACCGTTACTTCTGATCGAATATCTATCTGAG 419  
 DB 297 ATCTTCTGGAGCTCGAGGCTTACTACACCGTTACTTCTGATCGAATATCTATCTGAG 356  
 QY 420 CCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTTGGTGAAGCTGCTGCACCTGT 479  
 DB 357 CCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTTGGTGAAGCTGCTGCACCTGT 416  
 QY 480 CACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTGAGAA 539  
 DB 417 CACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTGAGAA 476

```

QY 540 CCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAAGAGAGATGAGCACTGCCATC 599
Db 477 CCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAAGAGAGATGAGCACTGCCATC 536
QY 600 TCCGACACCTCTCAGCAAGTTCAGGTGCGCATCAATAACAACCTCTATGTGCAACCACT 659
Db 537 TCCGACACCTCTCAGCAAGTTCAGGTGCGCATCAATAACAACCTCTATGTGCAACCACT 596
QY 660 CTTCTCAAGTACA-GTTTCCGCAAGGACATC-TTTGGAGACATGTTTGTGCTGGCAAT 717
Db 597 CTTCTCAAGTACAAGTTCCTCCGCAAGGACATCTTTGGAGACATGTTTGTGCTGGCAAG 656
QY 718 GCCCAAGCGGGAAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
Db 657 ATGCCAAGCGGCGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
QY 778 AATGG 782
Db 717 CAGG 721

RESULT 10
LOCUS BM018961
DEFINITION 603646918F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5428540 5',
mRNA sequence.
ACCESSION BM018961
VERSION BM018961.1 GI:16533315
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Description: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1896 row: c column: 05
High quality sequence start: 8
High quality sequence stop: 678.
Location/Qualifiers
1..824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5428540"
/clone_lib="NIH_MGC_98"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: brain; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley), using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 177 a 235 c 240 g 172 t
ORIGIN

Query Match 48.4%; Score 523.4; DB 13; Length 824;
Best Local Similarity 90.7%; Pred. No. 1.6e-124;
Matches 679; Conservative 0; Mismatches 36; Indels 34; Gaps 10;

```

```

QY 48 GCAGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCGCTGGAGTTCAGGAAGCC-GGAGT 106
Db 2 GCTGTGGCGCGCTGCTGCTGGCGCTGCTGCTGGCGCTGGAGTTCAGGAAGCC-TGGAGT 61
QY 107 CGCAGAGAGCGGGCGCTTATCAGGACCATGCGCGCGACGGGTATCATCATCGCGCATCG 166
Db 62 CGCAGAGAGCGGGCGCTTATCAGGACCATGCGCGCGACGGGTATCATCATCGCGCATCG 120
QY 167 TGGGTGGAGAGGAGCGCGCAACTCGGGCGCTGGCGCTGGCGAGGGAGCGCTGCGCTGTGG 226
Db 121 TGGGTGGAGAGGAGCGCGCAACTCGGGCGCTGGCGCTGGCGAGGGAGCGCTGCGCTGTGG 180
QY 227 ATTCCACAGTATCGGAGTGAGCTGCTCAGCCACCGCTGGCGACATCAC-GGCGGGCGCAC 285
Db 181 ATTCCACAGTATCGGAGTGAGCTGCTCAGCCACCGCTGGCGACATCACAGGGGGCGCAC 240
QY 286 -TGCTTTGAACCTATAGTGACCTTAGTGTATCCCTCCGGGTGATGGTCCAGTTTGCCCA 344
Db 241 ATGCTTTGAACCTATAGTGACCTTAGTGTATCCCTCCGGGTGATGGTCCAGTTTGCCCA 300
QY 345 GCTGAC-TTCCATGCCATCCTTCTGGAGCGCTGCAGCGCTTACTACCCCT-TACTTTCGTA 402
Db 301 GCTGACATTCATGCCATCCTTCTGGAGCGCTGCAGCGCTTACTACCCCTTACTTTCGTA 360
QY 403 TCGAATATCTATCTAGAGCCCTCGCTACCTGGGGAATTCACCTATGACATTCGCTTGGTG 462
Db 361 TCGAATATCTATCTAGAGCCCTCGCTACCTGGGGAATTCACCTATGACATTCGCTTGGTG 420
QY 463 AAGCTGTCTGCACCTGTACCTTACCTAACACATCCAGCCCATCTGTCTCAGGCGCTCC 522
Db 421 AAGCTGTCTGCACCTGTACCTTACCTAACACATCCAGCCCATCTGTCTCAGGCGCTCC 480
QY 523 ACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGGTGGGGGTACATCAAGAG 582
Db 481 ACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGGTGGGGGTACATCAAGAG 540
QY 583 GATGAGGCACTGCCATCTCCCAACACCTCCAGGAAGTTTCAGGTGCCATCATCAAA-CAA 641
Db 541 GATGAG-----GAAGGTACAGGTGCCATCAAAACCAA 575
QY 642 CTCATGTGCAACC-ACCTCTTCCTCAAGTACAGTTT-CCGCAAGGACATCTTTGGAGAC 599
Db 576 CTCATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCCGCAAGGACATCTATGGAGAC 635
QY 700 ATGGTTTGTGCTGGCAATGCCCAAGCGGGAAGGATGCTGCTTCGTTGCTCAGTACAGTGA 759
Db 636 ATGGTTTGTGCTGGCAATGCCCAAGCGGGAAGGATGCTGCTTCAGTACAGTGTCCATCA 695
QY 760 CCCTTGGCGCTGTAAACAAGATGGAGTGTG 788
Db 696 CACCACCTCACAGACCAGGAAGCATCCTG 724

RESULT 11
LOCUS AA1611187/c
DEFINITION z059b06.sl Stratagene pancreas (#937208) Homo sapiens cDNA clone
IMAGE:591155 3' similar to TR:E218488 E218488 TRYPTASE ;, mRNA
sequence.
ACCESSION AA1611187
VERSION AA1611187.1 GI:1735441
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 591)
AUTHORS Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

```



```
|||||
Db 302 CCAGCTGACTTCCATCCCTCTCTGGAGCCTCGAGCCCTACTACACCCGTTACTTCGT 361
QY 402 ATCAATATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTGGT 461
Db 362 ATCAATATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTGGT 421
QY 462 GAAGCTGCTGCACTCTGACCTTACACTAAACACATCCAGCCCATCTGCTCCAGGCCCTC 521
Db 422 GAAGCTGCTGCACTCTGACCTTACACTAAACACATCCAGCCCATCTGCTCCAGGCCCTC 481
QY 522 CACATTTGAGTTTGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATC 576
Db 482 CACATTTGAGTTAGCAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTCTCCTC 536

RESULT 13
BI257649
LOCUS 602967571F1 NTH_MGC_12 Homo sapiens cDNA clone IMAGE:5107384 5',
DEFINITION mRNA sequence.
ACCESSION BI257649
VERSION BI257649
KEYWORDS BI257649.1 GI:148113233
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 825)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11259 row: m column: 17
High quality sequence stop: 753.
Location/Qualifiers
1..825
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5107384"
/tissue_type="cervical carcinoma cell line"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT 162 a 245 c 237 g 181 t
ORIGIN
Query Match 46.4%; Score 501.6; DB 13; Length 825;
Best Local Similarity 95.4%; Pred. No. 6.7e-119;
Matches 605; Conservative 0; Mismatches 14; Indels 15; Gaps 8;
QY 120 GCCCTTATCAGCACCATCGGCGGAGGTATCATACGTCGCGCATC-TGTGGGTGAGAGG 178
Db 119 GACCATCGAGGACCATCGGCGGAGGTATCATACGTCGCGCATC-TGTGGGTGAGAGG 178
QY 179 AC-GCCGAACCTGGCGGCTTGGCGGAGGAGCCCTGCGCTGGGATTCACACCTA 237
Db 179 ACTGCGGAACCTGGCGGCTTGGCGGAGGAGGAGCCCTGCGCTGGGATTCACACCTA 237
QY 238 TCCGAGTAGCCTGCTCAGCAACCGCTGGGCACTCAGCGGGCGGCACTGCTTTGAACC 297
Db 238 TCGCGAGTAGCCTGCTCAGCAACCGCTGGGCACTCAGCGGGCGGCACTGCTTTGAACC 297

us-09-787-844-l.rst
|||||
QY 298 TATAGTGACCTTAGTGATCCCTCCGGTGATGTCAGTTTGGCCAGCTGACTTCATG 357
Db 297 -----TGACCTTAGTGATCCCTCCGGTGATGTCAGTTTGGCCAGCTGACTTCATG 351
QY 358 CCATCTCTTGGAGCCTGCGAGGCTTACTACACCCGTTACTTCGTAATATCTATCTG 417
Db 352 CCATCTCTTGGAGCCTGCGAGGCTTACTACACCCGTTACTTCGTAATATCTATCTG 411
QY 418 AGCCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTGGTGAAGCTGTCGACCT 477
Db 412 AGCCCTCGCTACCTGGGGAATTCACCCCTATGACATTCGCTGGTGAAGCTGTCGACCT 471
QY 478 GTCACTTACACTAAACACATCCAGCCCATCTGCTCCAGCCCTCCACATTCAGTTGAG 537
Db 472 GTCACTTACACTAAACACATCCAGCCCATCTGCTCCAGCCCTCCACATTCAGTTGAG 531
QY 538 AACCCGACAGACTGCTGGGTGACTGGC-TGGGGGTACATCAAGAGGAT-GAGGCACTGC 595
Db 532 AACCCGACAGACTGCTGGGTGACTGGC-TGGGGGTACATCAAGAGGATGAGGCACTGC 591
QY 596 CATCTCCCCACACCTCCAGGAAGTTCAAGTTCAGGTCGCCATCATAAACAACCTATGTCACAC 655
Db 592 CATCTCCCCACACCTCCAGGAAGTTCAAGTTCAGGTCGCCATCATAAACAACCTATGTCACAC 651
QY 656 ACCTCTTCTCAAGTACAGTTT-CCGCAAGGACATCTTTGGAG--ACATGGTTTGTGCT 711
Db 652 ACCTCTTCTCAAGTACAGTTTTCGCGAAGGACATCTTTGGAGGACATGTTTGTGCT 711
QY 712 GGCATGCCCAAGCGGGAAGGATGCCCTGCTTCG 745
Db 712 GGCATGCCCAAGCGGGAAGGATGCCCTGCTTCG 745

RESULT 14
AK006271
LOCUS 1050 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:1700023E12:protease, serine, 21, full insert
sequence.
ACCESSION AK006271
VERSION AK006271.1 GI:12839279
KEYWORDS HTC; CAP trapper
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA,
clone:lib:RIKEN full-length enriched mouse cDNA library
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
REFERENCE
1 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
```



QY 696 AGACATGCTTTGCTGCGAATGCCCCAAGCGGGAGGATGCTGCTTGGTGACTCAGG 755  
|||||  
Db 682 AGACATGCTTTGCTGCGAATGCCCCAAGCGGGAGGATGCTGCTTGGTGACTCAGG 741  
|||||  
QY 756 TGGACCCCTTGCTGTAACAAAGATGAGTCTGTATCAGATTTGAGTCTGGG 815  
|||||  
Db 742 AGGACCCCTTGCTGTAACAAAGATGAGTCTGTATCAGATTTGAGTCTGGG 801  
|||||  
QY 816 AGTGGGCTGTGCTGCGCCCAATCGCGCGGTCTACACCAATATCAGCCACTTTGA 875  
|||||  
Db 802 AATAGGCTGTGCTGCGCCCAATCGCGCGGTCTACACCAATATCAGCCACTTTGA 861  
|||||  
QY 876 GTGATCCAGAACTGATGCGCCAGAGTGCATGTCCAGCCAGACCCCTCCCTGGCCACT 935  
|||||  
Db 862 CTGATCCAGAACTGATGCGCCAGAGTGCATGTCCAGCCAGACCCCTCCCTGGCCACT 921  
|||||  
QY 936 ACTCTTTTCTCTCTCTGCGGCTCTCCACTCTCTGCGGCGGCTGTGAGCCCTACCTGAG 995  
|||||  
Db 922 ACTGTTTCTCTCTCTGCGGCTCTCCACTCTCTGCGGCGGCTGTGAGCCCTACCTGAG 981  
|||||  
QY 996 CCATCGAGCTGGGCGCCACTGCAAGTCAAGCGCCCTGCTCTCTCTCTGCTTTGGTA 1055  
|||||  
Db 982 TAGCTCACACCTGTG-----AGGTGAGGTGTCTCTCTCTCTCTCTCTCTCTCT 1032  
|||||  
QY 1056 ATAAACACATT 1066  
|||||  
Db 1033 ATAAACCTGTT 1043  
|||||

RESULT 15  
BM018436  
LOCUS 603646140P1 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5428064 5',  
DEFINITION mRNA sequence.  
ACCESSION BM018436  
VERSION BM018436.1 GI:16532790  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://mhc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LCM1894 row: 0 column: 09  
High quality sequence start: 3  
High quality sequence stop: 622.  
Location/Qualifiers  
1..624  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5428064"  
/clone\_lib="NIH\_MGC\_98"  
/tissue\_type="astrocytoma grade IV, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: POTB; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).

FEATURES  
source

BASE COUNT 123 a 183 c 185 g 133 t  
ORIGIN

Query Match 44.7%; Score 483.6; DB 13; Length 624;  
Best Local Similarity 97.9%; Pred. No. 2.8e-114;  
Matches 511; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 120 GCCCTTATAGGACCATCGGCCACGGGTGCATCACCTCGCGCATCGTGGTGAGAGGA 179  
|||  
Db 100 GACCATCCGAGGACCATCGGCCACGGGTGCATCACCTCGCGCATCGTGGTGAGAGGA 159  
|||  
QY 180 CGCCGAACCTCGGCGCTTGGCCGTGGCAGGAGGAGCTCGCGCTGTGGGATTCACACGATG 239  
|||  
Db 160 CGCCGAACCTCGGCGCTTGGCCGTGGCAGGAGGAGCTCGCGCTGTGGGATTCACACGATG 219  
|||  
QY 240 CGAGTAGGAGCTCTCAGCCACCGCTGGGACACTCAGCGGCGGCGCATCGCTTTGAAACCTA 279  
|||  
Db 220 CGAGTAGGAGCTCTCAGCCACCGCTGGGACACTCAGCGGCGGCGCATCGCTTTGAAACCTA 279  
|||  
QY 300 TAGTGACCTTAGTGATCCCTCGGGTGGATGATGATGATGATGATGATGATGATGATGAT 359  
|||  
Db 280 TAGTGACCTTAGTGATCCCTCGGGTGGATGATGATGATGATGATGATGATGATGATGAT 339  
|||  
QY 360 ATCTTCTGGAGCTCGAGGCTTACTACACCGCTTACTTCTGATGATGATGATGATGATGAT 419  
|||  
Db 340 ATCTTCTGGAGCTCGAGGCTTACTACACCGCTTACTTCTGATGATGATGATGATGATGAT 399  
|||  
QY 420 CCCTCGCTACCTGGGGAATTCACCTATGACATTCCTTGGTGAAGCTGTCTGACACCTGT 479  
|||  
Db 400 CCCTCGCTACCTGGGGAATTCACCTATGACATTCCTTGGTGAAGCTGTCTGACACCTGT 459  
|||  
QY 480 CACCTACACTAAACACATCCAGCCCATCTCTCTCCAGGCTCCACATTTGATTTGAGAA 539  
|||  
Db 460 CACCTACACTAAACACATCCAGCCCATCTCTCTCCAGGCTCCACATTTGATTTGAGAA 518  
|||  
QY 540 CCGGACAGACTGTGGGTGACTGGCTGGGGGTACATCAAGAGGATGAGGACACTGCCATC 599  
|||  
Db 519 CCGGACAGACTGTGGGTGACTGGCTGGGGGTACATCAAGAGGATGAGGACACTGCCATC 578  
|||  
QY 600 TCCCCACACCTCCAGGAAGTTTCCAGGTCGGCATCATATAACAA 641  
|||  
Db 579 TCCCCACA-CCTCCAGGAAGTTTCCAGGTCGGCATCATATAACAA 619  
|||

Search completed: May 10, 2003, 04:57:12  
Job time : 2084 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 02:50:58 ; Search time 351 Seconds  
(without alignments)  
6935.642 Million cell updates/sec

Title: US-09-787-844-1

Perfect score: 1081

Sequence: 1 agaggcagagggggcgctca.....acattccagttgatgcctgc 1081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	100.0	1081	20	AAx87151
2	1081	100.0	1081	21	AAA08505
3	1068.8	98.9	1087	21	AAAC2015
4	1063.2	98.4	1131	22	AAAS01698
5	1063.2	98.4	1131	22	AAAS01706
6	1055.8	97.7	1100	20	AAx52259
7	1055.8	97.7	1100	22	AAAF72417
8	1054.8	97.6	1077	24	AAAD27418
9	1052.4	97.4	1079	21	AAZ45673

10	1049.6	97.1	1100	19	AAV59119	Nucleotide sequenc
11	1048.8	97.0	1117	19	AAV59118	Nucleotide sequenc
12	1045.8	96.7	1082	20	AAx15337	Nucleic acid encod
13	1045.8	96.7	1082	24	AAAD27417	Human eosinophil s
14	1037.8	96.0	1082	22	AAAS02648	Human secreted pro
15	938.8	86.8	942	20	AAx15336	Nucleic acid encod
16	938.8	86.8	942	21	AAAG2005	Hydrophobic domain
17	752.6	69.6	1103	22	AAAF5271	Nucleotide sequenc
18	496.4	45.9	498	22	AAAS01708	WISH TADG-16 catal
19	494.8	45.8	498	22	AAAS01707	WISH TADG-16 catal
20	493.4	45.6	529	24	AAAS61822	Lung small cell ca
21	479	44.3	959	19	AAV59132	Nucleotide sequenc
22	475	43.9	918	22	AAI67198	Nucleotide sequenc
23	465.4	43.1	575	22	AAAS24506	Human ovarian PCR-
24	465.4	43.1	575	22	AAH83114	Human ovarian tumo
25	422.2	39.1	888	24	ABK31769	DNA encoding novel
26	413.4	38.2	1018	24	AAAD27415	Human eosinophil s
27	334.2	30.9	3866	19	AAV59133	Genomic sequence o
28	329.8	30.5	837	24	AAAD27419	Human eosinophil s
29	318.2	29.4	373	24	ABL80716	Human ovarian canc
30	165	15.3	1110	22	AAAF76994	Human protease T c
31	165	15.3	1129	22	AAAS21354	Human cDNA sequenc
32	165	15.3	1151	22	AAAD08286	Human secreted pro
33	165	15.3	1703	21	AAZ52473	Human secreted pro
34	163.2	15.1	980	19	AAV59136	Nucleotide sequenc
35	161.8	15.0	1212	22	AAI58544	Human polynucleoti
36	161.6	14.9	2457	24	ABK31795	DNA encoding novel
37	157.2	14.5	1157	22	AAI60330	Human polynucleoti
38	155.8	14.4	1130	22	AAAF77000	Fusion gene of pro
39	151.8	14.0	1322	21	AAAS61704	cDNA encoding mous
40	150.6	13.9	2662	24	AAAD3854	Human protease PRT
41	150.6	13.9	2810	22	AAK94500	Human full-length
42	149	13.8	1958	24	ABK12891	Human protease PRT
43	145.2	13.4	1122	24	AAAD37036	Mouse tryptase cDN
44	145	13.4	437	24	ABL81643	Human ovarian canc
45	133.4	12.3	393	24	AAAD27416	Human eosinophil s

#### ALIGNMENTS

RESULT 1  
ID AAX87151 standard; cDNA; 1081 BP.  
XX  
AC AAX87151;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE Human protease HUPM-3 cDNA.  
XX  
KW Protease; human; HUPM-3; cell proliferation; cancer;  
KW immune disorder; inflammation; therapy; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 40..984  
FT sig\_peptide /tag= a  
FT /tag= b  
FT /note= "putative signal peptide sequence"  
FT mat\_peptide 77..981  
FT /tag= b

XX  
FN WO9936550-A2.  
XX  
PD 22-JUL-1999.  
XX  
PF 12-JAN-1999; 99WO-US00655.  
XX  
PR 16-JAN-1998; 98US-0008271.  
XX





PF testicular cancers by measuring expression of a cancer specific gene -  
XX Claim 6; Page 31; 36pp; English.

XX The present sequence encodes a human cancer-specific gene (CSG),  
CC designated Prol04. The inventive methods comprise measuring expression  
CC of CSG in order to diagnose, stage, image and treat gynecologic and  
CC testicular cancers. The levels of CSG are compared to levels found in a  
CC normal human control (a change in the measured level of CSG is associated  
CC with the presence of the cancer). The early diagnosis of cancers  
CC improves the success rate of therapeutic protocols.

XX Sequence 1081 BP; 202 A; 321 C; 321 G; 237 T; 0 other;

Query Match 100.0%; Score 1081; DB 21; Length 1081;  
Best Local Similarity 100.0%; Pred. No. 1e-252;  
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGAGAGGGGGCTCAGGCGCGGGAGAGAGGCGCATGGGCGCGGGGGCGGTG 60  
DB 1 AGGAGGAGAGGGGGCGTCAAGCGCGGGAGAGAGGCGCATGGGCGCGGGGGCGGTG 60  
QY 61 CTGCTGGCGCTGCTGGCTGGGCTGGGCTGAGTCAAGAGCGGAGTGCAGGAGCGGCG 120  
DB 61 CTGCTGGCGCTGCTGGCTGGGCTGGGCTGAGTCAAGAGCGGAGTGCAGGAGCGGCG 120  
QY 121 CCTTATCAGACCATCGCGCGGACGGGTATCATCGTCGCGCATCGTGGGTGGAGAGGAC 180  
DB 121 CCTTATCAGACCATCGCGCGGACGGGTATCATCGTCGCGCATCGTGGGTGGAGAGGAC 180  
QY 181 GCCAACTCGGGCGTTGGCGGTGCGAGGGAGCGCTGCGGCTGTGGGATGCCAGTATGC 240  
DB 181 GCCAACTCGGGCGTTGGCGGTGCGAGGGAGCGCTGCGGCTGTGGGATGCCAGTATGC 240  
QY 241 GGAGTGGAGCTGCTCAGGACCGCTGGGCACTCAGCGGGCGGCACTGTTTGAACCTAT 300  
DB 241 GGAGTGGAGCTGCTCAGGACCGCTGGGCACTCAGCGGGCGGCACTGTTTGAACCTAT 300  
QY 301 AGTGACCTTGTGTATCCTCGGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA 360  
DB 301 AGTGACCTTGTGTATCCTCGGGGTGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCA 360  
QY 361 TCCCTTGGAGCTGCGAGGCGCTATACACCGGTTACTTCTGATCGAATATCTATCTAGC 420  
DB 361 TCCCTTGGAGCTGCGAGGCGCTATACACCGGTTACTTCTGATCGAATATCTATCTAGC 420  
QY 421 CCTCGTACCTGGGGAATTCACCTATGATGCTTGGTGAAGCTGCTGCACTGCTC 480  
DB 421 CCTCGTACCTGGGGAATTCACCTATGATGCTTGGTGAAGCTGCTGCACTGCTC 480  
QY 481 ACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCGCTCCACATTTGAGTTGAGAAC 540  
DB 481 ACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCGCTCCACATTTGAGTTGAGAAC 540  
QY 541 CGGACAGACTGCTGGGTGACTGGGTGGGGGTACATCAAGAGGATGAGGCACTGCCATCT 600  
DB 541 CGGACAGACTGCTGGGTGACTGGGTGGGGGTACATCAAGAGGATGAGGCACTGCCATCT 600  
QY 601 CCCACACCTCCAGGAGTTGAGTGGCCATCAATCAACTATATGTGCAACACCTC 660  
DB 601 CCCACACCTCCAGGAGTTGAGTGGCCATCAATCAACTATATGTGCAACACCTC 660  
QY 661 TTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGGTTTGGCTGGCAATGCC 720  
DB 661 TTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGGTTTGGCTGGCAATGCC 720  
QY 721 CAAGCGGGAGGATGCTGCTTCCGTTGACTCAGTGGAGCCCTTGGCTGTAAACAAGAT 780  
DB 721 CAAGCGGGAGGATGCTGCTTCCGTTGACTCAGTGGAGCCCTTGGCTGTAAACAAGAT 780  
QY 781 GGAGTGGTATCAGATTCAGTGGTGGAGTGGGAGTGGGCTGTGGTGGCCCAATCGG 840  
DB 781 GGAGTGGTATCAGATTCAGTGGTGGAGTGGGAGTGGGCTGTGGTGGCCCAATCGG 840

QY 841 CCCGTGTCTACACCAATATCAGCCACACCTTTGAGTGGATCCAGAGCTGATGCCCGAG 900  
DB 841 CCCGTGTCTACACCAATATCAGCCACACCTTTGAGTGGATCCAGAGCTGATGCCCGAG 900  
QY 901 AGTGCAATGCTCCAGCAGAGACCCCTCCCTGGCCACTACTCTTTTCCCTCTTCTCTGGGCT 960  
DB 901 AGTGCAATGCTCCAGCAGAGACCCCTCCCTGGCCACTACTCTTTTCCCTCTTCTCTGGGCT 960  
QY 961 CTCGCCACTCTGGGGCGGCTGTGAGCCTACCTGAGCCATCGAGCCATCGAGCCTGGGGCACTGCCA 1020  
DB 961 CTCGCCACTCTGGGGCGGCTGTGAGCCTACCTGAGCCTACCTGAGCCATCGAGCCTGGGGCACTGCCA 1020  
QY 1021 AGTCAGGCGCTGGTCTCTTCTGCTCTTGGTAAATAACACATTCACGTTGATGCCCTG 1080  
DB 1021 AGTCAGGCGCTGGTCTCTTCTGCTCTTGGTAAATAACACATTCACGTTGATGCCCTG 1080  
QY 1081 C 1081  
DB 1081 C 1081

## RESULT 3

AAA62015  
ID AAA62015 standard; DNA; 1087 BP.

XX AAA62015;

XX 02-FEB-2001 (first entry)

XX Hydrophobic domain protein cDNA HP03116 isolated from KB cells.

DE Human; secreted protein; membrane protein; hydrophobic domain;  
KW proliferation control; differentiation induction; material transport;  
KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;  
KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;  
KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
KW gene therapy; ss.

XX Homo sapiens.

XX WO200029448-A2.

XX 25-MAY-2000.

XX 17-NOV-1999; 99WO-JP06412.

XX 17-NOV-1998; 98JP-0326255.

XX 22-DEC-1998; 98JP-0364315.

XX 16-MAR-1999; 99JP-0069811.

XX 27-APR-1999; 99JP-0119299.

XX 19-MAY-1999; 99JP-0138169.

XX (SAGA ) SAGAMI CHEM RES CENT.

XX (PROT-) PROTEGENE INC.

XX Kato S, Kimura T;

XX WPI; 2000-387753/33.

XX P-PSDB; AAB12132.

XX Proteins comprising hydrophobic regions, such as secretory and membrane

XX proteins, useful in research and diagnostics and having various  
XX activities e.g. immunomodulatory, antiinflammatory, chemokinetic,  
XX hemostatic, thrombolytic -  
XX Claim 4; Page 274-276; 410pp; English.  
XX Secretory proteins play important roles in the proliferation control, the  
XX differentiation induction, the material transport and the biophylaxis of  
XX cells. Membrane proteins have important roles as signal receptors, ion  
XX channels and transporters. The present sequence is the coding sequence



SQ Sequence 1131 BP; 217 A; 332 C; 340 G; 242 T; 0 other;

**D**

QY 1020	AAGTCAGGCCCCCTGGTTCCTCCTCGATCTTGTTGGTAATAAACACATCCAGTTGATGCCCT	1079
<b>D</b> 1045	AAGTCAGGCCCCCTGGTTCCTCCTCGATCTTGTTGGTAATAAACACATCCAGTTGATGCCCT	1104

QY 1 AGGAGCAGAGGGGGCGTCAGGCCCGGGAGAGGAGGCCATGGCGCGCGCGGGGGCGCTG 60





PR 15-SEP-1999; 99WO-US21547.  
 PR 05-OCT-1999; 99WO-US23089.  
 PR 29-NOV-1999; 99WO-US28214.  
 PR 30-NOV-1999; 99WO-US28313.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 05-JAN-2000; 99WO-US00219.  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX  
 PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;  
 PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;  
 PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kijavini IJ;  
 PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;  
 PI Williams PM, Wood WI;  
 XX  
 DR WPI: 2001-081051/09.  
 DR P-PSDB; AAB80256.  
 XX  
 PT Sixty one nucleic acids encoding PRO polypeptides which are useful in  
 PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung  
 PT squamous cell carcinoma) and neurodegenerative diseases (e.g.  
 PT Alzheimer's disease) -  
 XX  
 PS Claim 2; Fig 91; 393pp; English.  
 XX  
 CC The present sequence is one of sixty one nucleic acids encoding novel  
 CC secreted and transmembrane PRO polypeptides. The PRO polypeptides are  
 CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung  
 CC squamous cell carcinoma), gastrointestinal disorders (e.g.  
 CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.  
 CC ischaemia, atherosclerosis), inflammatory disorders such as coronary  
 CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and  
 CC diabetes and retinal disorders such as retinitis pigmentosa.  
 CC The PRO nucleic acids have applications in molecular biology, including  
 CC use as hybridization probes, and in chromosome and gene mapping.  
 XX  
 SQ Sequence 1100 BP; 225 A; 321 C; 314 G; 240 T; 0 other;  
 Query Match 97.7%; Score 1055.8; DB 22; Length 1100;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-246;  
 Matches 1057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 21 GGCGCGGGAGAGAGGCCATGGCGCGCGCGCGCGCTGCTGCTGGCGTGTGCTGGC 80  
 DB 1 GGCGCGGGAGAGAGGCCATGGCGCGCGCGCGCGCTGCTGCTGGCGTGTGCTGGC 60  
 QY 81 TCGGGTGGACTCAGGAAGCCGAGTGCAGAGAGCGCGCGCTTATCAGGACCATGCGG 140  
 DB 61 TCGGGTGGACTCAGGAAGCCGAGTGCAGAGAGCGCGCGCTTATCAGGACCATGCGG 120  
 QY 141 CCACGGGTATCATCGCTCGGCATCGTGGTGAGAGAGCGCGAACTCGGGCGTTGGCC 200  
 DB 121 CCACGGGTATCATCGCTCGGCATCGTGGTGAGAGAGCGCGAACTCGGGCGTTGGCC 180  
 QY 201 GTGGCAGGGAGCGCTCGGCTGTGGGATTCGCCAGTATCGGAGTGCAGCTCTCAGCCA 260  
 DB 181 GTGGCAGGGAGCGCTCGGCTGTGGGATTCGCCAGTATCGGAGTGCAGCTCTCAGCCA 240  
 QY 261 CGCTGGGCATCATCGCGCGCGCACTGCTTTGAAACCTATAGTACCTTAGTATCCCTC 320  
 DB 241 CGCTGGGCATCATCGCGCGCGCACTGCTTTGAAACCTATAGTACCTTAGTATCCCTC 300  
 QY 321 CGGGTGGATGGTCCAGTTGGCAGCTGACTTCCATGCCATCCTTCAGAGCTTCAGGC 380  
 DB 301 CGGGTGGATGGTCCAGTTGGCAGCTGACTTCCATGCCATCCTTCAGAGCTTCAGGC 360  
 QY 381 CTACTACACCCGTTACTTCTGATCAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 440  
 DB 361 CTACTACACCCGTTACTTCTGATCAATATCTATCTGAGCCCTCGCTACCTGGGGAATTC 420

QY 441 ACCTATGACATTCCTGGTGAAGCTGTCTGCACCTGTACCTACCTACCTAAACATCA 500  
 DB 421 ACCCTATGACATTCCTGGTGAAGCTGTCTGCACCTGTACCTACCTAAACATCA 480  
 QY 501 GCCCATCTCTCCAGGCTCCACATTTGAGTTTGAGAACCGAGAGTGTGGGTGAC 560  
 DB 481 GCCCATCTCTCCAGGCTCCACATTTGAGTTTGAGAACCGAGAGTGTGGGTGAC 540  
 QY 561 TGGCTGGGGTACATCAAGAGGATGAGGCACTGCATCTCCACACCTCCAGGAAT 620  
 DB 541 TGGCTGGGGTACATCAAGAGGATGAGGCACTGCATCTCCACACCTCCAGGAAT 600  
 QY 621 TCAGTCCCATCATPAAACAACTCTATGTGCAACCACTCTTCTTCAAGTACAGTTTCG 680  
 DB 601 TCAGTCCCATCATPAAACAACTCTATGTGCAACCACTCTTCTTCAAGTACAGTTTCG 660  
 QY 681 CAAGGACATCTTGGAGACATGGTTGTGCTGGCAATGCCAAGCGGGAAGGATGCCTG 740  
 DB 661 CAAGGACATCTTGGAGACATGGTTGTGCTGGCAATGCCAAGCGGGAAGGATGCCTG 720  
 QY 741 CTTGGTCACTCAGTGGAGCCCTTGGCTGTAAAGAAATGAGTGTGATCAGATTGG 800  
 DB 721 CTTGGTCACTCAGTGGAGCCCTTGGCTGTAAAGAAATGAGTGTGATCAGATTGG 780  
 QY 801 AGTCTGAGCTGGGAGTGGCTGTGCTGGCCCAATGCCGCGGTGTCTACACCAATAT 860  
 DB 781 AGTCTGAGCTGGGAGTGGCTGTGCTGGCCCAATGCCGCGGTGTCTACACCAATAT 840  
 QY 861 CAGCCACCATCTTGGAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGCCAGCCAGA 920  
 DB 841 CAGCCACCATCTTGGAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGCCAGCCAGA 900  
 QY 921 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGCGGT 980  
 DB 901 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGCGGT 960  
 QY 981 CTGAGCTACTCTGAGCCCATGAGCTGGGCGCACTGCCAAGTCAAGCCCTGTCTCTT 1040  
 DB 961 CTGAGCTACTCTGAGCCCATGAGCTGGGCGCACTGCCAAGTCAAGCCCTGTCTCTT 1020  
 QY 1041 CTGCTTTGTTGTTAATAACACATTCAGTTGATGCT 1079  
 DB 1021 CTGCTTTGTTGTTAATAACACATTCAGTTGATGCT 1059  
 RESULT 8  
 AAD27418  
 ID AAD27418 standard; DNA; 1077 BP.  
 XX  
 AC AAD27418;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX  
 DE Human eosinophil serine protease-1 (esp-1) like enzyme. DNA #3.  
 KW Human; eosinophil serine protease-1; esp-1; enzyme; antiinflammatory;  
 KW antiasthmatic; antiallergic; osteopathic; cytostatic; dermatological;  
 KW asthma; allergy; chronic obstructive pulmonary disease; COPD;  
 KW osteoporosis; dermatitis; Paget's disease; therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200198503-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 20-JUN-2001; 2001WO-EP06936.  
 XX  
 PR 21-JUN-2000; 2000US-212844P.  
 PR 31-OCT-2000; 2000US-24417P.  
 PR 30-MAR-2001; 2001US-279766P.  
 XX









[illegible]

Query Match 97.0%; Score 1048.8; DB 19; Length 11117;  
Best Local Similarity 99.8%; Pred. No. 6.e-245;  
Matches 1050; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 GCCGCGGAGAGAGAGCCATGGCGCGCGCGGCGCTGCTGCTGGCGCTGCTGCTGGCT 81  
DB 1 GCCGCGGAGAGAGAGCCATGGCGCGCGCGGCGCTGCTGCTGGCGCTGCTGCTGGCT 60  
QY 82 CGGGCTGGACTCAGGAAGCGGAGTCGAGGAGCGCGCCCTTATCAGGACCATGCGGC 141  
DB 61 CGGGCTGGACTCAGGAAGCGGAGTCGAGGAGCGCGCCCTTATCAGGACCATGCGGC 120  
QY 142 CGACGGTTCATCAGCTCCGATCGTGGTGGAGAGACCGCAACTCGGCGTTGGCGG 201  
DB 121 CGACGGTTCATCAGCTCCGATCGTGGTGGAGAGACCGCAACTCGGCGTTGGCGG 180  
QY 202 TGGCAGGGAGCTCGCGCTGTGGATTCGCCAGTATCGGAGTGGAGCTGCTCAGCCAC 261  
DB 181 TGGCAGGGAGCTCGCGCTGTGGATTCGCCAGTATCGGAGTGGAGCTGCTCAGCCAC 240  
QY 262 CGCTGGGACATCAGGGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCC 321  
DB 241 CGCTGGGACATCAGGGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCC 300  
QY 322 GGGTGGATGTCAGTTTGGCCAGCTGACTTCCATGCGCATCTTCTGGAGCCTGCGGCC 381  
DB 301 GGGTGGATGTCAGTTTGGCCAGCTGACTTCCATGCGCATCTTCTGGAGCCTGCGGCC 360  
QY 382 TACTACACCCGTTACTTCTGATCGAATATCTATCTGAGCCCTCGCTACCTGGGAAATCA 441  
DB 361 TACTACACCCGTTACTTCTGATCGAATATCTATCTGAGCCCTCGCTACCTGGGAAATCA 420  
QY 442 CCCTATGACATGCTTGGTGAAGTGTCTGCACTGTGCACTACCTAACACATCCAG 501  
DB 421 CCCTATGACATGCTTGGTGAAGTGTCTGCACTGTGCACTACCTAACACATCCAG 480  
QY 502 CCATCTGCTCCAGGCTCCCATTTGAGTTGAGAACCGGACAGCTGCTGGTGAAT 561  
DB 481 CCATCTGCTCCAGGCTCCCATTTGAGTTGAGAACCGGACAGCTGCTGGTGAAT 540  
QY 562 GGTGGGGTACATCAAGAGGATGAGGACTGCCATCTCCACACCTCCAGGAGTT 621  
DB 541 GGTGGGGTACATCAAGAGGATGAGGACTGCCATCTCCACACCTCCAGGAGTT 600  
QY 622 CAGGTGCGCATATCAACAACTATGTGCAACCACTCTTCTCAAGTACAGTTTCCGC 681  
DB 601 CAGGTGCGCATATCAACAACTATGTGCAACCACTCTTCTCAAGTACAGTTTCCGC 660  
QY 682 AAGGACATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGCGGGAGGATGCCTGC 741  
DB 661 AAGGACATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGCGGGAGGATGCCTGC 720  
QY 742 TTCGTGACTCAGGTGGACCTTGGCTGCTTAAAGATGAGTGTGTTATCAGATTGA 801  
DB 721 TTCGTGACTCAGGTGGACCTTGGCTGCTTAAAGATGAGTGTGTTATCAGATTGA 780  
QY 802 GTCTGAGCTGGGAGTGGGCTGTGGTGGGCCAATCGCCCGGTGCTACACCAATATC 861  
DB 781 GTCTGAGCTGGGAGTGGGCTGTGGTGGGCCAATCGCCCGGTGCTACACCAATATC 840  
QY 862 AGCCACCACTTTGAGTGGATCCAGAGCTGATGGCCCAAGATGGCATGTCACAGCCAGAC 921  
DB 841 AGCCACCACTTTGAGTGGATCCAGAGCTGATGGCCCAAGATGGCATGTCACAGCCAGAC 900  
QY 922 CCCTCTGCGCCACTACTCTTTTCCCTCTCTGCTGGGCTCTCCACATCTCGGGCGGCTC 981  
DB 901 CCCTCTGCGCCACTACTCTTTTCCCTCTCTGCTGGGCTCTCCACATCTCGGGCGGCTC 960  
QY 982 TGAGCTTACCTGAGCCCATGAGCCTTGGGGCACTGCGCAAGTCAAGGCCCTGTTCTCTTC 1041  
DB 961 TGAGCTTACCTGAGCCCATGAGCCTTGGGGCACTGCGCAAGTCAAGGCCCTGTTCTCTTC 1020

QY 1042 TGCTCTGTTGGTAAATAAACACATTCAGTTG 1073  
DB 1021 TGCTCTGTTGGTAAATAAACACATTCAGTTG 1052

RESULT 12  
AA15337  
ID AA15337 standard; cDNA to mRNA; 1082 BP.  
XX AA15337;  
AC  
XX  
DT 04-MAY-1999 (first entry)  
XX  
DE Nucleic acid encoding a human eosinophil serine protease.  
XX Human; eosinophil; serine protease; allergic disease; infectious disease;  
KW tumour; granulomatous disease; collagen disease; vascular inflammation;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
PN JP11032768-A.  
XX  
PD 09-FEB-1999.  
XX  
PF 16-JUL-1997; 97JP-0191319.  
XX  
PR 16-JUL-1997; 97JP-0191319.  
XX  
PA (ONOI ) ONO PHARM CO LTD.  
XX  
DR WPI; 1999-183825/16.  
XX  
PT New eosinophils serine protease - useful for prevention and  
PT treatment of allergic, infectious, tumour, granulomatous and collagen  
PT diseases  
XX  
PS Claim 5; Page 10-11; 18pp; Japanese.  
XX  
CC The present sequence encodes a human eosinophil serine protease.  
CC The protease is useful in drug compositions for the prevention and  
CC treatment of allergic diseases, infectious diseases, tumour diseases,  
CC granulomatous diseases, collagen diseases and vascular inflammation.  
XX  
SQ Sequence 1082 BP; 216 A; 318 C; 308 G; 240 T; 0 other;

Query Match 96.7%; Score 1045.8; DB 20; Length 1082;  
Best Local Similarity 99.8%; Pred. No. 3.5e-244;  
Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 31 GAGGAGGCATGGCGCGCGCGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGGCTGA 90  
DB 1 GAGGAGGCATGGCGCGCGCGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGGCTGA 60  
QY 91 CTCAGGAAGCGGAGTTCGAGGAGCGCGGCGCTTATCAGGACCATGCGCGGCTG 150  
DB 61 CTCAGGAAGCGGAGTTCGAGGAGCGCGGCGCTTATCAGGACCATGCGCGGCTG 120  
QY 151 ATCAGCTGCGCATCTGCTGGTGGAGAGGACCGCAACTCGGCGCTTGGCGCTG 210  
DB 121 ATCAGCTGCGCATCTGCTGGTGGAGAGGACCGCAACTCGGCGCTTGGCGCTG 180  
QY 211 AGCTGCGCTGTGGGATTCACAGTATGCGGAGTGGAGCTGCTCAGCAGCGCTG 270  
DB 181 AGCTGCGCTGTGGGATTCACAGTATGCGGAGTGGAGCTGCTCAGCAGCGCTG 240  
QY 271 CTCAGGCGGCGGCTGCTTTGAACCTATAGTACCTTAGTATCCCTCGGCTGGAT 330  
DB 241 CTCAGGCGGCGGCTGCTTTGAACCTATAGTACCTTAGTATCCCTCGGCTGGAT 300  
QY 331 GTCAGTTTGGCGGCTGCTTCCATGCCATCTTCTGAGGCTGCGAGGCTTACTAC 390  
DB 301 GTCAGTTTGGCGGCTGCTTCCATGCCATCTTCTGAGGCTGCGAGGCTTACTAC 360

QY 391 CGTACTTCGATCGAATATATCTATCTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 450  
 DB 361 CGTACTTCGATCGAATATATCTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 420  
 QY 451 ATTGCTTGGTGAAGCTGTCTGCACTGTACCTACCTACCTAAACACATCCAGCCCATCTGT 510  
 DB 421 ATTGCTTGGTGAAGCTGTCTGCACTGTACCTACCTACCTAAACACATCCAGCCCATCTGC 480  
 QY 511 CTCAGGCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGG 570  
 DB 481 CTCAGGCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGG 540  
 QY 571 TACATCAAGAGGATGAGGCACTGCATCTCCACACCCCTCCAGGAAGTTTCAGGTGCC 630  
 DB 541 TACATCAAGAGGATGAGGCACTGCATCTCCACACCCCTCCAGGAAGTTTCAGGTGCC 600  
 QY 631 ATCATAAACAACCTATGTGTGAACACACCTCTTCCTCAAGTACAGTTTCCGCAAGSACATC 690  
 DB 601 ATCATAAACAACCTATGTGTGAACACACCTCTTCCTCAAGTACAGTTTCCGCAAGSACATC 660  
 QY 691 TTTGGAGACATGGTTGTGCTGCAATGCCCAAGCGGAGGATGCTGCTCGGTGAC 750  
 DB 661 TTTGGAGACATGGTTGTGCTGCAATGCCCAAGCGGAGGATGCTGCTCGGTGAC 720  
 QY 751 TCAGGTGGACCTTGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCTGTGAGC 810  
 DB 721 TCAGGTGGACCTTGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCTGTGAGC 780  
 QY 811 TGGGAGTGGGCTGTGCTGGCCCAATCGGCCCGGTGTCTACACATATCAGCCACAC 870  
 DB 781 TGGGAGTGGGCTGTGCTGGCCCAATCGGCCCGGTGTCTACACATATCAGCCACAC 840  
 QY 871 TTTGAGTGGATCCAGAGCTGATGCCCGCAGAGTGGCATGCCAGCCAGACCCCTCCCTGG 930  
 DB 841 TTTGAGTGGATCCAGAGCTGATGCCCGCAGAGTGGCATGCCAGCCAGACCCCTCCCTGG 900  
 QY 931 CCAGTACTCTTTTCCCTCTTCTTGGGCTCTCCACCTCTCCAGTGGGCGCGGTCTGAGCCCTAC 990  
 DB 901 CCAGTACTCTTTTCCCTCTTCTTGGGCTCTCCACCTCTCCAGTGGGCGCGGTCTGAGCCCTAC 960  
 QY 991 CTGAGCCCATGACGCTGGGCGCACTGCCAAGTCCAGGCTGGTCTCTTCTTCTGCTTGT 1050  
 DB 961 CTGAGCCCATGACGCTGGGCGCACTGCCAAGTCCAGGCTGGTCTCTTCTTCTGCTTGT 1020  
 QY 1051 TGGTAATAAACACATTCAGTTGATGCT 1079  
 DB 1021 TGGTAATAAACACATTCAGTTGATGCT 1049

RESULT 13  
 AAD27417  
 ID AAD27417 standard; DNA; 1082 BP.  
 XX  
 AC AAD27417;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX  
 DE Human eosinophil serine protease-1 (esp-1) like enzyme DNA #2.  
 XX  
 KW Human; eosinophil serine protease-1; esp-1; enzyme; antiinflammatory;  
 KW antihistaminic; antiallergic; osteopathic; cytostatic; dermatological;  
 KW asthma; allergy; chronic obstructive pulmonary disease; COPD;  
 KW osteoporosis; dermatitis; Paget's disease; therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200198503-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 20-JUN-2001; 2001WO-EP06936.  
 XX

21-JUN-2000; 2000US-212844P.  
 31-OCT-2000; 2000US-244171P.  
 30-MAR-2001; 2001US-279766P.  
 XX (FARB ) BAYER AG.  
 FA Xiao Y;  
 XX  
 PI  
 XX  
 DR WPI; 2002-122283/16.  
 XX  
 PT Novel purified human eosinophil serine protease 1-like enzyme, useful  
 PT for identifying modulators of enzyme activity for treating Paget's  
 PT disease, osteoporosis, airway allergy, asthma  
 XX  
 PS Disclosure; Fig 6; 131pp; English..  
 XX  
 CC The invention relates to a purified human eosinophil serine protease-1  
 CC (esp-1) like enzyme. Esp-1 like enzyme is useful in in-vitro or in-vivo  
 CC assays to identify test compounds with potential therapeutic or  
 CC diagnostic value. Esp1-like enzyme modulator is useful for treating esp1-  
 CC like enzyme dysfunction related diseases condition such as asthma, airway  
 CC allergy, chronic obstructive pulmonary disease (COPD) or osteoporosis.  
 CC Esp-1 like enzyme is also useful in diagnostic assays for detecting  
 CC diseases and abnormalities or susceptibility to diseases related to  
 CC presence of mutations in the nucleic acid sequences which encode the  
 CC enzyme. Pharmaceutical composition comprising esp-1 like enzyme is useful  
 CC for treating dermatitis, Paget's disease, and preventing degradation of  
 CC bone implants particularly dental implants. The present sequence is  
 CC human esp-1 like enzyme DNA.  
 XX  
 .SQ Sequence 1082 BP; 216 A; 318 C; 308 G; 240 T; 0 other;  
 Query Match 96.7%; Score 1045.8; DB 24; Length 1082;  
 Best Local Similarity 99.8%; Pred. No. 3.5e-244;  
 Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 31 GAGGAGGCCATGGGCGCGCGGGCGCTGTGCTGGCGCTGCTGCTGGCTGGCGCTGGA 90  
 DB 1 GAGGAGGCCATGGGCGCGCGGGCGCTGTGCTGGCGCTGCTGCTGGCTGGCGCTGGA 60  
 QY 91 CTCAGGAAGCCGAGTCCGAGGAGCGCGCGCTTATCAGGACCATCGCGCGGCGGCTC 150  
 DB 61 CTCAGGAAGCCGAGTCCGAGGAGCGCGCGCTTATCAGGACCATCGCGCGGCGGCTC 120  
 QY 151 ATCAGCTCGCGCATCGTGGTGGAGGAGCAGCGCAACTCGGCGCTTGGCGCTGGCAGGG 210  
 DB 121 ATCAGCTCGCGCATCGTGGTGGAGGAGCAGCGCAACTCGGCGCTTGGCGCTGGCAGGG 180  
 QY 211 AGCCTGCGCCTGTGGGATTCACGATATGCGGAGTGGAGCTGCTCAGCCACCGCTGGGCA 270  
 DB 181 AGCCTGCGCCTGTGGGATTCACGATATGCGGAGTGGAGCTGCTCAGCCACCGCTGGGCA 240  
 QY 271 CTCAGGCGGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCCGGGTGGATG 330  
 DB 241 CTCAGGCGGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCCGGGTGGATG 300  
 QY 331 GTCCAGTTTGGCCAGCTGACTTCCATGCGATCCCTTCTGGAGCTCGAGGCTACTACACC 390  
 DB 301 GTCCAGTTTGGCCAGCTGACTTCCATGCGATCCCTTCTGGAGCTCGAGGCTACTACACC 360  
 QY 391 CGTTACTTCGATCGAATATATCTATGTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 450  
 DB 361 CGTTACTTCGATCGAATATATCTATGTAGCCCTCGCTACCTGGGGAATTCACCCCTATGAC 420  
 QY 451 ATTGCTTGGTGAAGCTGTCTGCACTGTACCTACCTAAACACATCCAGCCCATCTGT 510  
 DB 421 ATTGCTTGGTGAAGCTGTCTGCACTGTACCTAAACACATCCAGCCCATCTGT 480  
 QY 511 CTCAGGCTCCACATTTGAGTTTGAAGAACCGGACAGACTGCTGGGTGACTGGCTGGGG 570  
 DB 481 CTCAGGCTCCACATTTGAGTTTGAAGAACCGGACAGACTGCTGGGTGACTGGCTGGGG 540  
 QY 571 TACATCAAGAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTTCAGGTGCC 630



Qy	599	CATGGTTTGTGCTGGCAATGCCCAAGCGGGGAAGGATGCTCTTCGGTGACTCAGGTGG	758
Db	669	CATGGTTTGTGCTGGCAATGCCCAAGCGGGGAAGGATGCTCTTCGGTGACTCAGGTGG	728
Qy	759	ACCTTTGGCCCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCTGACCTGGGGAGT	818
Db	729	ACCTTTGGCCCTGTAAACAAGATGGAGTGTGGTATCAGATTGGAGTCTGACCTGGGGAGT	788
Qy	819	GGGCTGTGGTGGCGCCCAATCGGCCCGGTGCTCTACACCAATATCAGCCACCACTTTTGAGTG	878
Db	789	GGGCTGTGGTGGCGCCCAATCGGCCCGGTGCTCTACACCAATATCAGCCACCACTTTTGAGTG	848
Qy	879	GATCCAGAAGCTGATGGCCCGCAGTGGCATGTCCAGCCAGACACCCCTCTCGGCCACTACT	938
Db	849	GATCCAGAAGCTGATGGCCCGCAGTGGCATGTCCAGCCAGACACCCCTCTCGGCCCTACT	908
Qy	939	CTTTTTCCTCCCTCTCTGGGCTCTCCCACTCTGGGGCGGCTCTGAGCCCTACCTTGAGCCCC	998
Db	909	CTTTTTCCTCCCTCTCTGGGCTCTCCCACTCTGGGGCGGCTCTGAGCCCTACCTTGAGCCCC	968
Qy	999	ATGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTCTCTCTCTCTGTGGTGAATA	1058
Db	969	ATGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTCTCTCTCTCTGTGGTGAATA	1028
Qy	1059	AACACATTCGATTGATGCTT	1079
Db	1029	AACACATTCGATTGATGCTT	1049

RESULT 15	
AAI15336	
ID	AAI15336 standard; cDNA to mRNA; 942 BP.
XX	
XX	AAI15336;
XX	
XX	04-MAY-1999 (first entry)
XX	
DE	Nucleic acid encoding a human eosinophil serine protease.
XX	
XX	Human; eosinophil; serine protease; allergic disease; infectious disease;
KW	tumour; granulomatous disease; collagen disease; vascular inflammation;
KW	ss.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FH	1..942
FT	/*tag= a
FT	/note= "no termination codon"
XX	
PN	JPI1032768-A.
XX	
PD	09-FEB-1999.
XX	
PF	16-JUL-1997; 97JP-0191319.
XX	
PR	16-JUL-1997; 97JP-0191319.
XX	
PA	(ONOY ) ONO PHARM CO LTD.
XX	
XX	WPI; 1999-183825/16.
DR	P-PSDB; RAAW7116.
XX	
PT	New eosinophils serine protease - useful for prevention and
PT	treatment of allergic, infectious, tumour, granulomatous and collagen
PT	diseases
XX	
PS	Claim 4; Page 10; 18pp; Japanese.
XX	
CC	The present sequence encodes a human eosinophil serine protease.
CC	The protease is useful in drug compositions for the prevention and
CC	treatment of allergic diseases, infectious diseases, tumour diseases,

CC	granulomatous diseases, collagen diseases and vascular inflammation.
XX	
SQ	Sequence 942 BP; 176 A; 283 C; 276 G; 207 T; 0 other;
	Query Match 86.88; Score 938.8; DB 20; Length 942;
	Best Local Similarity 99.88; Pred. No. 3.1e-218;
	Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps
QY	40 ATGGGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTCGGGCTGGACTCAGGAAG 99
DB	1 ATGGCGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTCGGGCTGGACTCAGGAAG 60
QY	100 CCGAGATCGCAGGAGCGGCGCCCTTATCAGGACCATTCAGGCGCGAGCGGTCATCAGCTCG 159
DB	61 CCGAGATCGCAGGAGCGGCGCGCTTATCAGGACCATTCAGGCGCGAGCGGTCATCAGCTCG 120
QY	160 CGCATCTGGGTGGAGAGGACGCCAACTCGGGCGTTGGCCGTGGCAGGGGAGCCTCGCG 219
DB	121 CGCATCTGGGTGGAGAGGACGCCAACTCGGGCGTTGGCCGTGGCAGGGGAGCCTCGCG 180
QY	220 CTGTGGGATTTCCACGATATGCGGAGTAGCCTGCTCAGCCACCGCTGGGCACCTCAGGGCG 279
DB	181 CTGTGGGATTTCCACGATATGCGGAGTAGCCTGCTCAGCCACCGCTGGGCACCTCAGGGCG 240
QY	280 GCGCACTGCTTTGAAACCTATAGTAGACCTTAGTAGATCCCTCCGGGTGGATGGTCCAGTTT 339
DB	241 GCGCACTGCTTTGAAACCTATAGTAGACCTTAGTAGATCCCTCCGGGTGGATGGTCCAGTTT 300
QY	340 GGCACGTGACTTCCATGCGCATCTTCTGGAGCCTGCGAGGCCTACTACACCCCTTACTTC 399
DB	301 GGCACGTGACTTCCATGCGCATCTTCTGGAGCCTGCGAGGCCTACTACACCCCTTACTTC 360
QY	400 GTATCGAATATCTATCTGAGCCCTCGCTACTCTGGGGAATTCACCCATATGACATTGCCTTG 459
DB	361 GTATCGAATATCTATCTGAGCCCTCGCTACTCTGGGGAATTCACCCATATGACATTGCCTTG 420
QY	460 GTGAAGCTGTCTGCACCTGTCACTACACTAAACACATCCAGGCCCATCTGTCTCCAGGCC 519
DB	421 GTGAAGCTGTCTGCACCTGTCACTACACTAAACACATCCAGGCCCATCTGTCCCTCAGGCC 480
QY	520 TCCACATTTGAGTTTGAGAACCGACACAGCTGCTGGGTGACTGGCTGGGGGTACATCAAA 579
DB	481 TCCACATTTGAGTTTGAGAACCGACACAGCTGCTGGGTGACTGGCTGGGGGTACATCAAA 540
QY	580 GAGGATGAGGCACCTGCCATCTCCACACCCCTCCAGGAAGTTTCAGGTGCGCATCAAAAC 639
DB	541 GAGGATGAGGCACCTGCCATCTCCACACCCCTCCAGGAAGTTTCAGGTGCGCATCAAAAC 600
QY	640 AACTCTATGTGCACACCACTCTTCCCTCAAGTACAGTTTCCGAAGGACATCTTTGGAGAC 699
DB	601 AACTCTATGTGCACACCACTCTTCCCTCAAGTACAGTTTCCGAAGGACATCTTTGGAGAC 660
QY	700 ATGTTTGTGCTGGCAATGCCAAGGCGGAAGGATGCCCTGCTTCGGTGACTCAGGTGGA 759
DB	661 ATGTTTGTGCTGGCAATGCCAAGGCGGAAGGATGCCCTGCTTCGGTGACTCAGGTGGA 720
QY	760 CCTTTGGCCTGTAAACAGAAATGGACTGTGTGTATCAGATTGGAGTGTGAGCTGGGAGTG 819
DB	721 CCTTTGGCCTGTAAACAGAAATGGACTGTGTGTATCAGATTGGAGTGTGAGCTGGGAGTG 780
QY	820 GGCTGTGCTGGGCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGG 879
DB	781 GGCTGTGCTGGGCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGG 840
QY	880 ATCCAGAAGCTGTATGGCCACAGTGGCATGTCCACAGCCAGACCCCTCCTGGCCACTACTC 939
DB	841 ATCCAGAAGCTGTATGGCCACAGTGGCATGTCCACAGCCAGACCCCTCCTGGCCACTACTC 900
QY	940 TTTTTCCTCTTCTCTGGGCTCTCCCACTCTCTGGGGCCGGTC 981
DB	901 TTTTTCCTCTTCTCTGGGCTCTCCCACTCTCTGGGGCCGGTC 942

901 TTTTTCCTCTTCTCTGGGCTCTCCCACTCCTGGGGCCGGTC 942

Mon May 12 16:37:38 2003

Search completed: May 10, 2003, 03:24:32  
Job time : 355 secs

us-09-787-844-1.rng

---

Page 16

STIC-Biotech/ChemLib

93059

From: Yu, Misook  
Sent: Friday, May 02, 2003 1:50 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/787,844

Please search SEQ ID NO:1.

Examiner Misook Yu, Ph.D.  
703-308-2454 (Phone)  
Art Unit 1642  
CM1-8E18 (Room)  
CM1-8E12 (Mail Box)

Mary Jane Ruhl  
Tech. Info. Specialist, STIC  
TC-1600  
CM-1, Room 6A-06  
Phone: 605-1155

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 5/2/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 03:17:32 ; Search time 75 Seconds  
(without alignments)  
4420.234 Million cell updates/sec

Title: US-09-787-844-1

Perfect score: 1081

Sequence: 1 agaggcagagggggcgctca.....acattccagtgtgctgc 1081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	100.0	1081	4	US-09-008-271A-15
2	752.6	69.6	1103	4	US-09-386-642-59
3	165	15.3	1110	4	US-09-386-653A-1
4	155.8	14.4	1130	4	US-09-386-653A-8
5	131.4	12.2	1430	4	US-09-386-629-1
6	126.8	11.7	1166	4	US-09-386-629-2
7	118.6	11.0	1108	2	US-09-016-366A-14
8	118.6	11.0	1108	2	US-08-978-404B-20
9	117.8	10.9	1142	4	US-09-386-642-8
10	117.8	10.9	1169	4	US-09-386-642-7
11	117	10.8	1137	2	US-09-016-366A-18
12	117	10.8	1137	2	US-08-978-404B-13
13	116.2	10.7	1783	4	US-09-510-738A-188
14	116.2	10.7	2363	4	US-09-742-703-3
15	115.4	10.7	735	4	US-09-079-970A-1
16	115.4	10.7	771	4	US-09-079-970A-4
17	115.4	10.7	1081	2	US-09-016-366A-22
18	115.4	10.7	1081	2	US-08-978-404B-17
19	115.4	10.7	1128	2	US-09-016-366A-20
20	115.4	10.7	1128	2	US-08-978-404B-15
21	115.4	10.7	1154	2	US-09-016-366A-16
22	115.4	10.7	1154	2	US-08-978-404B-11
23	115.4	10.7	1219	2	US-08-978-404B-7
24	112	10.4	2413	4	US-09-518-046-1
25	111.2	10.3	1605	2	US-09-000-846-1
26	106	9.8	1031	2	US-08-978-404B-1
27	105	9.7	654	4	US-09-280-116-57

28	104.2	9.6	1097	2	US-08-978-404B-4	Sequence 4, Appli
29	104	9.6	1155	4	US-09-163-951-15	Sequence 15, Appl
30	104	9.6	1240	4	US-09-163-951-14	Sequence 14, Appl
31	102.6	9.5	1103	2	US-09-016-366A-24	Sequence 24, Appl
32	100	9.3	1109	4	US-09-088-651-6	Sequence 6, Appli
33	97	9.0	1109	4	US-09-088-651-1	Sequence 1, Appli
34	96	8.9	2416	4	US-09-261-416-1	Sequence 1, Appli
35	95.8	8.9	683	4	US-09-088-651-3	Sequence 3, Appli
36	93.8	8.7	1095	2	US-08-978-404B-9	Sequence 9, Appli
37	93.8	8.7	2544	4	US-09-518-046-3	Sequence 3, Appli
38	85.6	7.9	1440	1	US-07-882-202A-3	Sequence 3, Appli
39	85.6	7.9	1440	1	US-08-021-615A-3	Sequence 3, Appli
40	85.6	7.9	1440	1	US-08-321-777-3	Sequence 3, Appli
41	85.6	7.9	1440	3	US-09-009-217-13	Sequence 13, Appl
42	85.6	7.9	1440	3	US-09-009-656-13	Sequence 13, Appl
43	85.6	7.9	1440	5	PCT-US93-04493-3	Sequence 3, Appli
44	85.6	7.9	2422	1	US-08-475-845-1	Sequence 1, Appli
45	85.6	7.9	2422	1	US-08-327-690-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-008-271A-15  
; Sequence 15, Application US/09008271A  
; Patent No. 6203979  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; Hillman, Jennifer L.  
; Yue, Henry  
; Guegler, Karl J.  
; Corley, Neil C.  
; Tang, Tom Y.  
; Shan, Purvi  
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,271A  
FILING DATE: 16-Jan-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Mohan-Peterson, Sheela  
REGISTRATION NUMBER: 41,201  
REFERENCE/DOCKET NUMBER: PF-0458 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1081 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSTUT03  
CLONE: 789927  
SEQUENCE DESCRIPTION: SEQ ID NO: 15 :  
US-09-008-271A-15

Query Match									
Best Local Similarity 100.0%; Score 1081; DB 4; Length 1081;									
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps									
QY	1	AGAGGACAGAGGGGGCGTCAGGCCGCGGGAGAGGAGGCATCTGGCGCGCGGGGGCGCTG	60						
DB	1	AGAGGACAGAGGGGGCGTCAGGCCGCGGGAGAGGAGGCATCTGGCGCGCGGGGGCGCTG	60						
QY	61	CTGCTCGGCGCTGCTGCTGCGCTCGGCGCTGGACTCAGGAAGCCGGAGTCGAGAGGCGGCG	120						
DB	61	CTGCTCGGCGCTGCTGCTGCGCTCGGCGCTGGACTCAGGAAGCCGGAGTCGAGAGGCGGCG	120						
QY	121	CCCTTATCAGGACCATTGGCGGCGAGCGGTCATCAGCTCGGCATCTGTGGGTGGAGAGGAC	180						
DB	121	CCCTTATCAGGACCATTGGCGGCGAGCGGTCATCAGCTCGGCATCTGTGGGTGGAGAGGAC	180						
QY	181	GCGGAACCTGGCGGTTGGCCGTGGCAGGGGAGCCTGGCCCTGTGGGATTCGCCAGTATGC	240						
DB	181	GCGGAACCTGGCGGTTGGCCGTGGCAGGGGAGCCTGGCCCTGTGGGATTCGCCAGTATGC	240						
QY	241	GGAGTGAGCCCTGCTCAGGCACCGCTGGGCACCTCAGCGGGGCGACCTGCTTTGAAACCTAT	300						
DB	241	GGAGTGAGCCCTGCTCAGGCACCGCTGGGCACCTCAGCGGGGCGACCTGCTTTGAAACCTAT	300						
QY	301	AGTGACCTTAGTGATCCCTCCGGGTGGATGGTCAGCTTTGGCCAGCTGACTTCCATGCCA	360						
DB	301	AGTGACCTTAGTGATCCCTCCGGGTGGATGGTCAGCTTTGGCCAGCTGACTTCCATGCCA	360						
QY	361	TCCCTCTGGAGCCTGCAGGCCCTACTACACCGGTTACTTCGATTCGAAATATCTATCTGAGC	420						
DB	361	TCCCTCTGGAGCCTGCAGGCCCTACTACACCGGTTACTTCGATTCGAAATATCTATCTGAGC	420						
QY	421	CCTCGCTACCTTGGGGAATTCACCCATATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTC	480						
DB	421	CCTCGCTACCTTGGGGAATTCACCCATATGACATTCGCTTGGTGAAGCTGTCTGCACCTGTC	480						
QY	481	ACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTCAGTTTGAGAAC	540						
DB	481	ACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCTCCACATTCAGTTTGAGAAC	540						
QY	541	CGGACAGACTGCTGGGTGACTGGCTGGGGTGACATCAAGAGATGAGGCATGCGCATCT	600						
DB	541	CGGACAGACTGCTGGGTGACTGGCTGGGGTGACATCAAGAGATGAGGCATGCGCATCT	600						
QY	601	CCCCACACCCCTCCAGGAAGTTCAAGTTCGCCATCATAAACAACTCTATGTGCAACCACTC	660						
DB	601	CCCCACACCCCTCCAGGAAGTTCAAGTTCGCCATCATAAACAACTCTATGTGCAACCACTC	660						
QY	661	TTCTCTAAGTACAGTTTTCGCAAGGACATCTTTGGAGACATGGTTGTCTGGCAATGCC	720						
DB	661	TTCTCTAAGTACAGTTTTCGCAAGGACATCTTTGGAGACATGGTTGTCTGGCAATGCC	720						
QY	721	CAAGGGGGGAAGATGCGCTTCGGTGACTCAGGTGGACCCCTGGCCCTGTAAACAAGAT	780						
DB	721	CAAGGGGGGAAGATGCGCTTCGGTGACTCAGGTGGACCCCTGGCCCTGTAAACAAGAT	780						
QY	781	GGACTGTGCTATCAGATTGGAGTCGTGAGCTGGGAGTGGGCTGTGGTGGCCCAATCGG	840						
DB	781	GGACTGTGCTATCAGATTGGAGTCGTGAGCTGGGAGTGGGCTGTGGTGGCCCAATCGG	840						
QY	841	CCCGGTGTCTACACCAATATCAGCACCCACTTTGAGTGGATCCAGAAGCTGATGCCCCAG	900						
DB	841	CCCGGTGTCTACACCAATATCAGCACCCACTTTGAGTGGATCCAGAAGCTGATGCCCCAG	900						
QY	901	AGTGGCATCTCCAGCCAGACCCCTCTTGGCCACTACTCTTTTCCCTCTCTCTGGGCT	960						
DB	901	AGTGGCATCTCCAGCCAGACCCCTCTTGGCCACTACTCTTTTCCCTCTCTCTGGGCT	960						
QY	961	CTCCCACTCCTGGGGCGGCTGTGAGCCCTACCTGAGCCCATCGACGCTGGGGCCACTGCCA	1020						
DB	961	CTCCCACTCCTGGGGCGGCTGTGAGCCCTACCTGAGCCCATCGACGCTGGGGCCACTGCCA	1020						

```

QY   1021 AGTCAGGCCCGGTTCTCTTCTGTCTGGTAAATAAACACATTC CAGTTGATGCCTG 1080
      |||
Db    1021 AGTCAGGCCCGGTTCTCTTCTGTCTGGTAAATAAACACATTC CAGTTGATGCCTG 1080
      |||

QY   1081 C 1081
      |
Db    1081 C 1081

RESULT 2
US-09-386-642-59
; Sequence 59, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jenson
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
; OTHER INFORMATION: sequence of human protease F in CFEX2 zymogen
; OTHER INFORMATION: vector
US-09-386-642-59

```

Query Match	69.6%;	Score 752.6;	DB 4;	Length 1103;
Best Local Similarity	93.6%;	Pred. No. 5.1e-186;		
Matches 785;	Conservative 0;	Mismatches 54;	Indels 0;	Gaps 0;
QY 163	ATCGTGGGTGGAGAGACCGCCGAAC	TGCGGCGTTGGCCGTGGCAGGGAGCCTGGCCCTG	222	
Db				
QY 139	ATCGTTGGGGGTATGCTCTAGAAC	TGCGGCGTTGGCCGTGGCAGGGAGCCTGGCCCTG	198	
Db				
QY 223	TGGGATTCCACGTATGCGGAGT	GAGCCTGCTCAGCCACCGCTGGGCACTCACGGCGGG	282	
Db				
QY 199	TGGGATTCCACGTATGCGGAGT	GAGCCTGCTCAGCCACCGCTGGGCACTCACGGCGGG	258	
Db				
QY 283	CACCTGCTTGAACCTATAGTGAC	CTTAGTATCCCTCCGGTGGATGGTTCAGTTTGGC	342	
Db				
QY 259	CACCTGCTTGAACCTATAGTGAC	CTTAGTATCCCTCCGGTGGATGGTTCAGTTTGGC	318	
Db				
QY 343	CAGCTGACTTCCATGCCATCCTTC	TGAGAGCCTGAGGCGTACTACACCGGTACTTCGTA	402	
Db				
QY 319	CAGCTGACTTCCATGCCATCCTTC	TGAGAGCCTGAGGCGTACTACACCGGTACTTCGTA	378	
Db				
QY 403	TCGAATATCTATCTGAGCCCTCG	CTTACCTGGGGAATTCACCCCTATGACATTCGCTTGGTG	462	
Db				
QY 379	TCGAATATCTATCTGAGCCCTCG	TACCTGGGGAATTCACCCCTATGACATTCGCTTGGTG	438	
Db				
QY 463	AAGCTGCTGCACCTGTACACCTAC	ACTAACAACATCCAGCCCATCTGTCCAGGCCCTCC	522	
Db				
QY 439	AAGCTGCTGCACCTGTACACCTAC	ACTAACAACATCCAGCCCATCTGTCTCCAGGCCCTCC	498	
Db				
QY 523	ACATTGAGTTTGAGAACCGGACAG	ACTGCTGGGTGACTGGCTGGGGGTACATCAAGAG	582	
Db				
QY 499	ACATTGAGTTTGAGAACCGGACAG	ACTGCTGGGTGACTGGCTGGGGGTACATCAAGAG	558	
Db				
QY 583	GATGAGGACTTGCCATCTCCCCAC	ACCCTCCAGGAAGTTCAGGTGCGCATCATAAACAAC	642	
Db				
QY 559	GATGAGGACTTGCCATCTCCCCAC	ACCCTCCAGGAAGTTCAGGTGCGCATCATAAACAAC	618	
Db				
QY 643	TCTATGTGCAACCACTCTTCCCTCA	GTACAGTTTCGCAAGGACATCTTTGGAGACATG	702	
Db				
QY 619	TCTATGTGCAACCACTCTTCCCTCA	GTACAGTTTCGCAAGGACATCTTTGGAGACATG	678	
Db				

```
QY 703 GTTTGTGCTGCGCAATGCCAAGCGGGAAGATGCCCTGCTCGGTGACTCAGGTGGACCC 762
Db 679 GTTTGTGCTGCGCAATGCCAAGCGGGAAGATGCCCTGCTCGGTGACTCAGGTGGACCC 738
QY 763 TTGGCCCTGACAGAAATGACTGTGCTATCAGATTGGAGTCTGAGCTGGGAGTGGCC 822
Db 739 TTGGCCCTGACAGAAATGACTGTGCTATCAGATTGGAGTCTGAGCTGGGAGTGGCC 798
QY 823 TGTGCTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATC 882
Db 799 TGTGCTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATC 858
QY 883 CAGAACTGATGGCCAGAGTGGCATGTCCAGCCAGCCCTCTCTGCCCACTACTCTTT 942
Db 859 CAGAACTGATGGCCAGAGTGGCATGTCCAGCCAGCCCTCTCTGCCCACTACTCTTT 918
QY 943 TTCCCTCTCTCTGCTGCTCCCACTCTCTGGGCCGCTCTGAGCTACCTGAGCCCATG 1001
Db 919 CATCACCATCACTAGCGGCCGCTTCCCTTTAGTGAGGGTTAAATGCTTCGAGCAGACATG 977

RESULT 3
US-09-386-653A-1
; Sequence 1, Application US/09386653A
; Patent No. 6458564
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/09/386, 653A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-386-653A-1

Query Match 15.3%; Score 165; DB 4; Length 1110;
Best Local Similarity 53.7%; Pred. No. 1.1e-33;
Matches 429; Conservative 0; Mismatches 340; Indels 30; Gaps 3;

QY 113 AGCGCGCCCTTATCAGGACCATGCGCGGAGCGGTATCATCGCTCGCGCATCGTGGTG 172
Db 88 AGAGGGCCAAAGGACGACACAGCCCTGTGCTGCGCCCGAGGATGCTGAACCGAATGTTGGCG 147
QY 173 GAGAGACGCGCAACTCGGCGGTGGCGGTGGCAGGAGGAGCTGCGCTGTGGGATTCCT 232
Db 148 GCGAGACGCGGAGGCGGAGTGGCCCTGGCAAGTACAGATCCAGCCGCAACGGAAGCC 207
QY 233 ACCTATGCGGAGTGGCTCTCAGCCAGCGCTGGGCACTCAGCGCGGCGCACTGCTTTG 292
Db 208 ACTTCTCGGGGCGAGCTCATCGCGAGCAGTGGTCTTGCAGCGCTGCGCACTGCTTCC 267
QY 293 AAACCTATAGTACCTTAGTATGATCCCTCCCGGTGGATGCTCAGTTGGCCAGCTGACTT 352
Db 268 GCAACACCTCTGAGC-----GTCCCTGTACACAGGCTCCTGTGGGGCAA 312
QY 353 CCATGCCATCTTCTGAGGCTCGAGGCTTACTACACCGCTTACTTCTGATCAATATCT 412
Db 313 GCGAGTATGAGCGCGGACACACCGCTATGATGCGCGGTGAGGAGGAGCA 372
QY 413 ATCTGAGCCCTCGCTACCTGGGAATTCACCCCTATGACATTTGCTTGGTGAAGTCTGTG 472
Db 373 ACC---CCCTGTACAGGCGACGCGCTCCAGCGCTGAGCTGGCCCTGTGGAGCTGGAGG 429
QY 473 CACCTGTACCTACATAACACATCCAGGCCATCTCTCTCAGGCGCTCCACATTTAGT 532
Db 430 CACAGTGGCCCTTACACCAATATACATCTCCCGGTGTGCTGCTGCTGACCCCTCGGTGATCT 489
```

```
QY 533 TTGAGAACCGACAGACTGCTGGTGTGCTGGGGTACATCAAGAGGATGAGGCAC 592
Db 490 TTGAGACGGGCATGAACCTGCTGGTCTACTGGCTGGGCGAGCCAGTGAGGAACACCTCC 549
QY 593 TGGCATCTCCCAACACCCCTCCAGGAAGTTCCAGTTCGCGCATCAATAAACAATCTATGTGCA 652
Db 550 TGGCCGAACCGCGGATCTGTCAGAACTCGCTGTGCCATCATCGACACACCAAGTGCA 609
QY 653 ACCACCTCTTCCCTCAAGTACA-----GTTTCCGCAAGGACATCTTTGGAGACA 700
Db 610 ACCTGCTCTACAGCAAGACACACCGAGTTTGGCTACCAACCCCAACCATCAAGATGACA 669
QY 701 TGGTTTGTGCTGGCAATGCCCAAGCGGGAAGGATGCTTCTCGGTGACTCAGGTGGAC 760
Db 670 TGTGTGCGCGGCTTCGAGGAGGCAAGAGGATGCTGCAAGGCGGACTCGGCGGCGC 729
QY 761 CTTGGGCTGTAAACAAGATGGACTGTGGTATCAGATTGGAGTCTGAGCTGGGAGTGG 820
Db 730 CCCTGGTGTGCTGCTGGGTGAGTGGTGTGAGGCGGGGTGATCAGCTGGGGTGGG 789
QY 821 GCTGTGTCGCGCCCAATCGGCCGCTGTCTACACCAATATCAGCCACCACTTTGAGTGG 880
Db 790 GCTGTGCGCGCAGAACCGCCAGGTGTCTATCTCCGTGTACCGCCACCAACATGGA 849
QY 881 TCCAGAAGCTGATGGCCCA 899
Db 850 TCCATCGGATCATCCCA 868

RESULT 4
US-09-386-653A-8
; Sequence 8, Application US/09386653A
; Patent No. 6458564
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/09/386, 653A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of
; OTHER INFORMATION: Protease T in a zymogen activation vector
US-09-386-653A-8

Query Match 14.4%; Score 155.8; DB 4; Length 1130;
Best Local Similarity 60.4%; Pred. No. 2.6e-31;
Matches 281; Conservative 0; Mismatches 172; Indels 12; Gaps 1;

QY 447 TGACATTCCTGTTGGTGAAGCTGTGCACTGTGCTACCTACACTAAACATCCAGCCAT 506
Db 432 TGAGTGGCCCTGGTGGAGCTGGAGGCCAGCCAGTGCCTTCCACCAATTACATCTCCCGT 491
QY 507 CTGTCTCCAGGCTCCACATTTGAGTTTGAACCGGACAGACAGTCTGGGTGACTGGCTG 566
Db 492 GTGCTCTGCTGACCCCTCGGTGATCTTTGAGACGGGATGAACCTGCTGGGTCACTGGCTG 551
QY 567 GGGGTACATCAAGAGGATGAGGCATCTCCCATCTCCCAACACACCTCCAGGAAGTTCAGGT 626
Db 552 GGGCAGCCCCAGTGAAGAGACCTCTTCCCGAACCAGGATCTCTGCAGAAATCAGCTGT 611
QY 627 CGCATCATAAACAATCTATGTGCAACCACTCTTCTCTCAAGTACA-----G 674
Db 612 GCCCATCATGACACACCCCAAGTCAACCTGCTCTACAGCAAGACACCCAGGTTTGGCTA 671
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1166
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-E catalytic
; OTHER INFORMATION: domain in a zymogen activated construct
US-09-386-629-2

Query Match
Best Local Similarity 11.7%; Score 126.8; DB 4; Length 1166;
Matches 363; Conservative 0; Mismatches 357; Indels 12; Gaps 1;

QY 163 ATCGTGGGTGGAGAGACGCGAACTCGGGCGTTGGCCGTGGCAGGGGAGCGCTGCGCGTG 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 ATCGTGGGGGCTATGCTCTAGAGGACGAGCGTGGCCCTGGATCGTGCATCCAGAAG 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 223 TGGGAATCCACGATATCGGAGTGAAGCCTGCTCAGCCACCGCTGGGCACACACGGCGCG 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 AATGGGACCCACCACCTGCGCAGGTTCTGCTCACCAGCCGCTGGGTGATCACTGCTGCC 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 283 CACTGCTTTGAACACTATAGTGACCTTAGTAPCCCTCCGGGTGGATGTGCCAGTTTGGC 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 CACTGTTTCAAGGACAACTGAAACAAACCATACCTGTTCTGTCTGTCTGTGGGGGCGCTGG 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 343 CAGCTGACTTCCATGCATCCTTCTGGAGCCTCGAGGCCCTACTACACCGTTACTTCGTA 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 CAGCTGGGGAACCTGGCTCT-----CGTCCCAGAAGGTGGGTGTGCTG 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 403 TCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCATATGACATTTGCTTTGGTG 462
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 GTGGAGCCCCACCTGTGTATTCCTTGAAGGAAGTGCTGTGTGCAGACATTCGCCCTGGTG 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 463 AAGCTGTCTGCACCTGTACCTACACTAAACACATCCAGCCCATCTGTCTCCAGCCCTCC 522
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 454 CGTCTCGAGCGCTCATACAGTTCTCAGACGGGGTCTCCCATCTGCTACCTGATGCC 513  
QY 523 ACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGTACATCAAGAG 582  
Db 514 TCTATCCACCTCCCTCCAAACACCCACCTGCTGGATCTCAGGCTGGGGAGCATCCAAGAT 573  
QY 583 GATGAGGACACTGCCATCTCCACACACCCTCCAGGAAGTTCAGGTGCGCATATAAACAC 642  
Db 574 GGAGTTCCCTTGGCCCAACCTCAGACCCCTGCAGAGAGCTGAAGTTCTTATCATGACATCG 633  
QY 643 TCTATGTCGACCACTCTTCTCCTCAAGTACAGTTTCCGACAGGACATCTTTGGAGACATG 702  
Db 634 GAAGTCTCCAGCCATCTCTACTTGGCGGGAGCAGACAGGACCCACTCAGTGAAGCATG 693  
QY 703 GTTCTGCTGGCAATGCCAAGCGGGAGGATGCTCTCGGTGACTCAGGTGACCC 762  
Db 694 CTGTGTCGGGTACTTGGAGGGGAGGGGATGCTTCTGGGGGATCTCCGGGGGCCCC 753  
QY 763 TTGGCCTCTAACAAGAAATGGACTGTGTGATCAGATTGGAGTCTGAGTGGGGAGTGGGC 822  
Db 754 CTCATGTCCAGGTGGACGGCGCTGCTGCTGGCGGCATCATCAGCTGGGGCGAGGCG 813  
QY 823 TGTGTCGGCCCACTGCGCCGGTCTTACACCAATATCAGCCACCATTTGAGTGGATC 882  
Db 814 TGTGCCGAGCGCAACAGCCCGGGGTCTACATCAGCCCTCTGTGGCGCACCGCTCTGGGTG 873  
QY 883 CAGAGCTGATG 894  
Db 874 GAGAGATCGTG 885

## RESULT 7

US-09-016-366A-14  
; Sequence 14, Application US/09016366A  
; Patent No. 5955431

## ; GENERAL INFORMATION:

; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston

; STATE: MA

; COUNTRY: U.S.A.

; ZIP: 02210-2211

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/016,366A

; FILING DATE: January 30, 1998

; CLASSIFICATION: 530

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/037,090

; FILING DATE: 05-FEB-1997

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Plumer, Elizabeth R.

; REGISTRATION NUMBER: 36,637

; REFERENCE/DOCKET NUMBER: B0801/7093

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-720-3500

; TELEFAX: 617-720-2441

## ; TELEX:

; INFORMATION FOR SEQ ID NO: 14:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 1108 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

## ; TOPOLOGY: linear

## ; MOLECULE TYPE: cDNA

US-09-016-366A-14

## Query Match

11.0%; Score 118.6; DB 2; Length 1108;

Best Local Similarity 56.0%; Pred. No. 1.1e-21;

Matches 248; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

QY 448 GACATTGGCTTGGTGAAGCTGTCTGCACCTGTCACTTACACTAAACACATCCAGCCCATC 507  
Db 397 GACGTTGCCCTGCTGGAGCTTGGAGTCCCTGTGAATGTCTCCACCATATCCACCCCATATA 456  
QY 508 TGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGTGGGTGACTGGCTGG 567  
Db 457 TCCCTGCCCTCGCTCGGAGACCTTCCCTCCCTGGACATCGTGTGGGTGACAGCTGG 516  
QY 568 GGTGATACAAAGAGATGAGGCACTGCCATCTCCCCACACCTCCAGGAAGTTCAGGTTC 627  
Db 517 GCGGACATTGATAATGACGAGCCTCTCCACCTCTTATCTCTGAAGCAAGTGAAGTT 576  
QY 628 GCCATCAATAACAACACTCTATGTGCAACCACTCTTCTCAAGTACAGTTTCCGCAAGGAC 687  
Db 577 CCCATGTGGAACACAGCCTGTGTGACCGGAAGTACACACTGGCTCTACACGGGAGAT 636  
QY 688 ATCTTTGGAGACATGGTTTGTG-----CTGGCAATGCCCAAGCGGGAAGGATGCCCTGC 741  
Db 637 GATTTTCCCATTTGCTCATGATGGCATGCTGTGTGCTGGAATATACAGGAGAGACTCCTGC 696  
QY 742 TTCGGTGACTCAGGTGGGACCTTGGCCCTGTACAGAAATGGAATGGACTGTGGTATCAGATTGGA 801  
Db 697 CAGGCGGATTCAGGGGGGCCACTGCTGTGCAAAAGTGAAGGTTACCTGGCTGCAGCAGGA 756  
QY 802 GTCGTGAGCTGGGGAGTGGCTGTGGTGGCCCAATCGCCCGGTGTCTACACCAATATC 861  
Db 757 GTGTGACCTGGGGTGGAGGCTGCGCACAGCCCAACAGCCCTGGCATCTACACCCGGGTG 816  
QY 862 AGCCACCACTTTGAGTGGATCCA 884  
Db 817 ACATACTACTTAGACTGGATCCA 839

## RESULT 8

US-08-978-404B-20

; Sequence 20, Application US/08978404B

; Patent No. 5968782

## ; GENERAL INFORMATION:

; APPLICANT: Stevens, Richard L.

; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES

; TITLE OF INVENTION: FIBRINOGEN

; NUMBER OF SEQUENCES: 74

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

; STREET: 600 Atlantic Avenue

; CITY: Boston

; STATE: MA

; COUNTRY: U.S.A.

; ZIP: 02210-2211

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

## ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/978,404B

; FILING DATE: 25-NOV-97

; CLASSIFICATION: 435

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/032,354

; FILING DATE: 04-DEC-1996

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Plumer, Elizabeth R.

; REGISTRATION NUMBER: 36,637

; REFERENCE/DOCKET NUMBER: B0801/7090



Db 534 CTGCACTGTCTACTGGTGGGTCTATGTGGCCCTCAGTGAGCCTCTCTGAGCCCAAGCC 593  
Qy 609 CTTCCAGGAAGTCTCAGGTGCGCATATAAACAACCTCTATGTGCAACCACTCTCTCTCAA 668  
Db 594 ACTGAGCAACTCAGGTGCGCTCTGATCAGTCGTGAGACGTGTACTCCCTGTACAACT 653  
Qy 669 GTACAGTTTCGCAAGGA-----CATCTTTGGAGACATGTTTGTGCTGGCAATGC 719  
Db 654 CGAGGCCAAGCTCAGGAGCGCACTTTGTCCAAAGAGGACATGTTGTGCTGGCTATGT 713  
Qy 720 CCAGGGGGGAAGTACCTCTCTCGGTGACTCAGGTGAGCCCTGGCTGTAAACAAGAA 779  
Db 714 GGAGGGGGGAAGAGCGCTCCAGGGTGTACTCTGGGGCCCACTCTCTCTCCCTGTGGA 773  
Qy 780 TGGAGTGTGGTATCAGATTTGAGTCTGAGTGGGAGTGGGTGTGTCGGCCCAATCG 839  
Db 774 GGGTCTCTGTACTCAGCGGCAATGTGAGCTGGGAGATCCCTGTGGGGCCCGCAAG 833  
Qy 840 GCGCGGTCTACACCAATATACGCCACCACTTTGAGTGTGATCCAGAGCTGATGGC 896  
Db 834 GCCTGGTGTATACACTCTGGCCTCCAGCTATGCCTCTCTGGATCCAAAGAGGTGAC 890

RESULT 11  
US-09-016-366A-18  
; Sequence 18, Application US/09016366A  
; Patent No. 5955431  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette.  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,366A  
; FILING DATE: January 30, 1998  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/037,090  
; FILING DATE: 05-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7093  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1137 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-016-366A-18

Query Match 10.88; Score 117; DB 2; Length 1137;  
Best Local Similarity 55.88; Pred. No. 3e-21;  
Matches 247; Conservative 0; Mismatches 190; Indels 6; Gaps 1;

Qy 448 GACATTTGCTTGGTGAAGCTGTCTGCACTGTCTCACTTAAACACATCCAGCCCATC 507  
Db 357 GACATCGCCCTGCTGGAGCTGAGGAGCGGTGACGTCTCCAGGCCACGTCCACACGGTC 416  
Qy 508 TGTCTCCAGGCTCCACATTTGAGTTTGAACCCGACAGACTGCTGGGTGACTGGCTGG 567  
Db 417 ACCGTGCCCCCTGCTCAGAGACCTTCCCGGGGATGCCGTGCTGGGTCACTGGCTGG 476  
Qy 568 GGGTACATCAAGAGGATGAGGCACTGCTATCTCCCAACACCTCCAGGAAGTTTCAGTGC 627  
Db 477 GCGATGTGGACATGATGAGGCGCTCCACCGCCATTTCTCTGAACAGAGTGAAGTGC 536  
Qy 628 GCCATATTAACAACCTCTATGTG-----CAACACACCTCTTCTCTCAAGTACAGTTTCGCG 681  
Db 537 CCATATATGGAACACCATTTGTGACGCAAAATACCACCTTGGCGCTCACACGGGAGAC 596  
Qy 682 AAGNACATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGGCGGGAAGATCCCTGCC 741  
Db 597 GACGTCCCATCGTCCGTCGACACATGCTGTGTCGGGGAACACCCGAGGGGACTCATGC 656  
Qy 742 TTCGTGACTCAGGTGGACCTTGGCCCTGTAAACAAGATGGACTGTGTATCAGATTGGA 801  
Db 657 CAGGCGACTCCGAGGCGCCCTGGTGTGCAAGGTGAATGGACCTGGCTGGAGCGGCG 716  
Qy 802 GTCGTGAGCTGGGAGTGGGTGTGTCGGCCCAATCGCCCGGTGTCTTACACCAATATC 861  
Db 717 GTGTGACGTGGGCGAGGGTGTGCCAGCCCAACCGGCTGGCATCTACACCGGTGTC 776  
Qy 862 AGCCACCACTTTGAGTGGATCCA 884  
Db 777 ACCTACTACTTGGATGGATCCA 799

RESULT 12  
US-08-978-404B-13  
; Sequence 13, Application US/08978404B  
; Patent No. 5968782  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES  
; TITLE OF INVENTION: FIBRINOGEN  
; NUMBER OF SEQUENCES: 74  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/978,404B  
; FILING DATE: 25-NOV-97  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/032,354  
; FILING DATE: 04-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7090  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1137 base pairs  
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-978-404B-13

Query Match
Best Local Similarity 10.8%; Score 117; DB 2; Length 1137;
Matches 247; Conservative 0; Mismatches 190; Indels 6; Gaps 1;

QY 448 GACATTGCGCTTGTGAGCTGTGACCTGTGACCTACATTAACACATCCAGCCCATC 507
DB 357 GACATGCGCCCTGTGAGCTGTGAGGAGCGGTGAACGCTTCCAGCCACGTCACACGGTC 416
QY 508 TGTCTCCAGGCCCTTACAGACCTTCCCGCGGGGATCCCGTGTGGTCACTGGCTGG 476
DB 417 ACCTGCGCCCTTCCAGACCTTCCCGCGGGGATCCCGTGTGGTCACTGGCTGG 476
QY 568 GGTATACATCAAGAGAGGAGGACCTGCGCATCTCCACACACCTCCAGGAATTCAGTGC 527
DB 477 GCGATGTGCAATGATGAGCGCCTCCACCGCCATTTCCCTCTGAAGCAGGTGAAGTGC 536
QY 628 GCATCAFAAACAACCTATATG-----CAACACCTCTTCTCTCAAGTACAGTTCCGC 681
DB 537 CCCATAATGGAACACCATTTGTGACGCAAAATACCACCTTGGCGCTACACGGGAGAC 596
QY 682 AAGGACATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGGCGGGGAGGATCCCTGC 741
DB 597 GAGTCCCGATCGTCCGTGACGACATGCTGTGCGCGGGAACACCGCGAGGACTCATGC 656
QY 742 TTCGTGACTCAGTGGAGCCCTTGGCCCTGTAAACAAGATGGACTGTGCTATCAGATTGA 801
DB 657 CAGGCGAGCTCGGAGGCGCCCTGGTGTGCAAGGTGAATGGACCTGCGTGCAGCGGGC 716
QY 802 CTCGTGACTGGGAGTGGCTGTGTCGGCCCAATCGGCCCGGTGTCTACACCAATATC 861
DB 717 GTGGTGTGCTGGGAGGCTGTGCGCAGCCCAACCGGCTGCGATCTACACCGCTGC 776
QY 862 AGCCACCACTTTGAGTGGATCCA 884
DB 777 ACCTACTTGTGACTGGATCCA 799

RESULT 13
US-09-510-738A-188
; Sequence 188, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 09/039,211
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 188
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length cDNA of hepsin
US-09-510-738A-188

Query Match
Best Local Similarity 10.7%; Score 116.2; DB 4; Length 1783;
Matches 398; Conservative 0; Mismatches 353; Indels 30; Gaps 4;

QY 114 GCGCGCGCTTATCAGGACCATCGCGCGGCTATCAGTCCGATCGTGGTGG 173
DB 683 GCGCGCGCTTATCAGGACCATCGCGCGGCTATCAGTCCGATCGTGGTGG 173
QY 174 AGAGACCGCACTTCGGCGCTTGGCGGTGGCAGGGAGCTGCGCTGTGGATCCCA 233
DB 174 AGAGACCGCACTTCGGCGCTTGGCGGTGGCAGGGAGCTGCGCTGTGGATCCCA 233

; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-742-703-3

Query Match
Best Local Similarity 10.7%; Score 116.2; DB 4; Length 2363;
Matches 1444; Conservative 0; Mismatches 1324; Indels 49; Gaps 0;

QY 743 CCGGACACACGCTTGGCGCGGTGGCAAGTCAAGCTTCCGCTATGATGAGACACA 802
QY 234 CGTATGGGAGTGAAGCTGCTCAGCCACCGCTGGGACTCAGCGCGGCGGACTGCTTTGA 293
DB 803 CCTCTGTGGGGATCCTGCTCTCCGGGACTGGGTGCTCAGACCGCGGCGGCTT--- 859
QY 294 AACCTATGAGCTTGTAGTATCCCTCCGGGTGGATGGTCCAGTTTGGCCAGCTGACTTC 353
DB 860 CCGGAGCGGAACCGGCTCTCTCCCGATGGGAGTGTGGCGGTGGCGGCGGCGGCG 919
QY 354 CATGCCATCTCTTGTGAGCTGCGAGCTTACACCCGCTTACTTCTGATATGATATCTA 413
DB 920 CTCTCCCGAGCTGCTGAGCTGGGGTGGAGCTGGGTGCTACCGAGGGGGTATCTTCC 979
QY 414 TGTGAGCCCTTCCCTACCTGGGGAATACCTTATGACATTTGGCTTGGTGAAGCTGCTGC 473
DB 980 CTCTGGGAGCCCAACAGCGAGGAGAACAGCAACGATATGGCTTGGTCCACTCTCCAG 1039
QY 474 ACCTGTACCTACATTAACACATCCAGCCATCTGTCTCCAGGCTCCACATTTGAGTT 533
DB 1040 TCCCTGCGCTTACAGAAATACATCCAGCTGTGTCCCTCCAGCTGCCGCGGCGGCT 1099
QY 534 TGAGAACCGGACAGACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 593
DB 1100 GGTGATGGCAAGATCTGTACCGTGGGCTGGGGCAACAGCAGTACTATGGCCAA 1159
QY 594 GCAATCTCCACACCCCTCCAGGAGTTCAGGTGGCCATCAATAAACAATCTATGTGCA 653
DB 1160 G-----GCCGGGTACTTCCAGAGGCTGAGTCCCGCATATCAGCATGATGTCTGCA 1213
QY 654 CCACCTCTTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATGTTTGGTGG 713
DB 1214 TGGCGCTGACTTC-----TATGAAACCATGATCAAGCCCAAGATGTTCTGTGCTGG 1264
QY 714 CAATCCCAAGCGGAGGATGCTTCTCGGTGACTCAGTGGAGCTTGGCGCTGTA 773
DB 1265 CTACCCGAGGCTGATGCTGCTCCAGGCGAGCGGTGGTGGTGGTGGTGGTGGTGG 1324
QY 774 CAA-----GAATGAGTGTGTATCAGATTTGGAGTTCGTGAGTGGGAGTGGG 821
DB 1325 GGACAGCATCTCTGGAGCGGCTGGCGCTGTGGCATTTGTGAGTTGGGCGACTGG 1384
QY 822 CTGTGTCGCGCCCAATCGGCGGCTGTCTACCAATATCAGCCACCATTTGAGTGGAT 881
DB 1385 CTGTGTCGCGCCCAAGCGGCTGTACCAATATCAGCCACCATTTGAGTGGAT 1444
QY 882 C 882
DB 1445 C 1445

RESULT 14
US-09-742-703-3
; Sequence 3, Application US/09742703
; Patent No. 6423543
; GENERAL INFORMATION:
; APPLICANT: Patrick Allen Marcotte
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF HEPsin EXPRESSION
; FILE REFERENCE: RTS-0090
; CURRENT APPLICATION NUMBER: US/09/742,703
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (826)..(2079)
US-09-742-703-3

Query Match
10.7%; Score 116.2; DB 4; Length 2363;

```



Best Local Similarity 51.0%; Pred. No. 6e-21;  
Matches 398; Conservative 0; Mismatches 353; Indels 30; Gaps 4;

QY 114 GCGGCGCCCTTATCAGACCAATGCGCGGCGGAGGCTATCAGCTCGCGCATCTGGTGG 173  
Db 1263 GCGCGCCATCTGCCAAGACTGCGCGGAGGAGTGCCTGCGGAGGAGG 1322

QY 174 AGAGACGCGCAATCTGCGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 233  
Db 1323 CGGGACACAGCTTGGCGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1382

QY 234 CATTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 293  
Db 1383 CACTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1439

QY 294 AACCTATAGTACCTTAGTATCTCTCGGCTGCGGAGGAGGAGGAGGAGGAGGAGG 353  
Db 1440 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1499

QY 354 CATGCCATCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 413  
Db 1500 CTCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1559

QY 414 TCTGAGCGCTCTGCTACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 473  
Db 1560 CTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1619

QY 474 ACCTCTACCTACCTAACACATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 533  
Db 1620 TCCCTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1679

QY 534 TGAGAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 593  
Db 1680 GGTGATGCGAGATCTGTACCGTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1739

QY 594 GCGATCTCCGACACCTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 653  
Db 1740 G-----GCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1793

QY 654 CCACCTCTCTCAAGTACAGTTCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 713  
Db 1794 TGGCGCTGACTTC-----TATGAAACAGATCAAGCCCAAGATGCTGTGCTGG 1844

QY 714 CAATGCCCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 773  
Db 1845 CTACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1904

QY 774 CAA-----GAATGAGTCTGGTATCAGATTTGAGTCTGAGTCTGGGAGGAGGAGGAGG 821  
Db 1905 GGACAGCATCTCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1964

QY 822 CTGTGCTCGGCGCAATCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 881  
Db 1965 CTGTGCTCGGCGCAATCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2024

QY 882 C 882  
Db 2025 C 2025

RESULT 15  
US-09-079-970A-1  
; Sequence 1, Application US/09079970A  
; Patent No. 6274366  
; GENERAL INFORMATION:  
; APPLICANT: Maffitt, Mark A.  
; APPLICANT: Niles, Andrew L.  
; APPLICANT: Haak-Frendscho, Mary  
; TITLE OF INVENTION: Enzymatically-Active Recombinant Human  
; TITLE OF INVENTION: Beta-Tryptase and Method of Making Same  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Intellectual Property Department

STREET: 8000 Excelsior Drive, Suite 401  
CITY: Madison  
STATE: WISCONSIN  
COUNTRY: U.S.A.  
ZIP: 53717-1914  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,970A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Leone, Joseph T.  
REGISTRATION NUMBER: 37,170  
REFERENCE/DOCKET NUMBER: 34506.073  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 831-2100  
TELEFAX: (608) 831-2106  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..735  
US-09-079-970A-1

Query Match 10.7%; Score 115.4; DB 4; Length 735;  
Best Local Similarity 55.5%; Pred. No. 6.7e-21;  
Matches 246; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

QY 448 GACATTCGCTTGGTGAAGCTGTCTGCACCTGTGCACCTACACCTAAACACATCCAGGCCATC 507  
Db 271 GACATTCGCTTGGTGAAGCTGTCTGCACCTGTGCACCTACACCTAAACACATCCAGGCCATC 330

QY 508 TGCTCCAGGCTCCACATTTGAGTTGAGACCGGAGAGTGTGGTGGTGTGCTGCTGCTGCTGCTG 567  
Db 331 ACCCTGCGGCTTGGTGAAGCTGTCTGCACCTGTGCACCTACACCTAAACACATCCAGGCCATC 390

QY 568 GGGTACATCAAGAGAGTGGAGCTGTCTGCACCTGTGCACCTACACCTAAACACATCCAGGCCATC 627  
Db 391 GCGATGTGGACAATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGGTGAAGTGC 450

QY 628 GCCATCAATAACAATCTATGTG-----CAACACCTCTTCTCTCAAGTACAGTTTCCGC 681  
Db 451 CCCATAATGAAACACACATTTGTGACGCAAAATACACCTTTGGGCGCTTACACGGGAGAC 510

QY 682 AAGGACATCTTTGGAGACATGTTGTGTGCTGCAATGCCAAGGCGGGAGGAGTGCCTGC 741  
Db 511 GACGTCCGATCTCGTGCAGACATGCTGTGTGCGGGAGACACCGGAGGAGTGCATGC 570

QY 742 TTCGGTACTCAGGTGGACCTTGGCTTGAACAAGATGGACTGTGGTATCAGATTGA 801  
Db 571 CAGGCGGACTCCGAGGCGGCTGTGTGCAAGGTGAATGGACCTTGGTGTGAGGCGGCG 630

QY 802 GTCGTGAGTGGGAGTGGGCTGTGGTTCGCGCCCAATTCGCGCGGCTGTCTTACACCAATATC 861  
Db 631 GTGGTCAAGTGGGCGGAGGCTGTGCGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 690

QY 862 AGCCACCACTTTGAGTGGATCCA 884  
Db 691 ACCTACTTGGACTGGATCCA 713

Mon May 12 16:37:39 2003

us-09-787-844-1.rni

Page 10

Search completed: May 10, 2003, 04:22:18  
Job time : 80 secs

.....

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 10, 2003, 02:58:22 ; Search time 3359 Seconds  
(without alignments)  
9365.922 Million cell updates/sec

Title: US-09-787-844-1  
Perfect score: 1081  
Sequence: 1 agaggcagagggggcgtca.....acattccagttgatgctgc 1081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1081	100.0	1081	6	AR142617	AR142617 Sequence
2	1054.8	97.6	1077	6	AX370460	AX370460 Sequence
3	1054.8	97.6	1077	9	AF058300	AF058300 Homo sapi
4	1045.8	96.7	1082	6	AX370459	AX370459 Sequence
5	1045.8	96.7	1082	6	E21866	E21866 Novel acido
6	1045.8	96.7	1082	6	E21867	E21867 Novel acido
7	1045.8	96.7	1082	9	AB031329	AB031329 Homo sapi
8	1039.8	96.2	1071	9	AB031330	AB031330 Homo sapi
9	1034.4	95.7	1085	6	E21853	E21853 Novel acido
10	1034.4	95.7	1085	6	E21854	E21854 Novel acido
11	969.2	89.7	1119	9	AB031331	AB031331 Homo sapi
12	938.8	86.8	942	6	E21865	E21865 Novel acido
13	927.4	85.8	945	6	E21852	E21852 Novel acido
14	493.4	45.6	529	6	AX321357	AX321357 Sequence
15	489	45.2	1055	10	AY005145	AY005145 Mus muscu
16	489	45.2	1084	10	AB033822	AB033822 Mus muscu
17	489	45.2	1095	10	AF176209	AF176209 Mus muscu
18	487.8	45.1	1063	10	AB059415	AB059415 Mus muscu
C 19	465.4	43.1	575	6	AX198283	AX198283 Sequence
C 20	465.4	43.1	575	6	AX208847	AX208847 Sequence
21	422.2	39.1	888	6	AX360070	AX360070 Sequence
22	413.4	38.2	1018	6	AX370454	AX370454 Sequence
23	338.6	31.3	1128	10	AB049453	AB049453 Mus muscu
24	334.2	30.9	5153	9	AF058301	AF058301 Homo sapi
C 25	334.2	30.9	35677	9	AC005361	AC005361 Homo sapi
C 26	332.6	30.8	199706	2	AC097329	AC097329 Pan trogl
27	329.8	30.5	837	6	AX370461	AX370461 Sequence
C 28	300	27.8	211962	9	AC092117	AC092117 Homo sapi
29	166.4	15.4	5015	10	AF226710	AF226710 Mus muscu
30	166.4	15.4	5551	10	AF304012	AF304012 Mus muscu
31	166.4	15.4	5628	10	AB059414	AB059414 Mus muscu
32	166.4	15.4	17405	10	AB041645	AB041645 Mus muscu
33	165	15.3	1013	9	HSA306593	HSA306593 Homo sapi
34	165	15.3	1077	9	AY030095	AY030095 Homo sapi
35	165	15.3	1129	6	AX464088	AX464088 Sequence
36	164.8	15.2	205939	2	AC110262	AC110262 Mus muscu
37	161.6	14.9	2457	6	AX360096	AX360096 Sequence
38	161.6	14.9	2681	6	AX480935	AX480935 Sequence
39	160	14.8	1488	10	EC024903	EC024903 Mus muscu
40	155.8	14.4	2066	9	AK055576	AK055576 Homo sapi
41	153	14.2	1259	10	RRAS642	AJ005642 Rattus ra
42	150.6	13.9	1199	9	AF175522	AF175522 Homo sapi
43	150.6	13.9	2662	6	AX338508	AX338508 Sequence
44	149	13.8	1958	6	AX342632	AX342632 Sequence
45	148.2	13.7	199706	2	AC097329	AC097329 Pan trogl

ALIGNMENTS

RESULT 1  
AR142617  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AR142617  
Sequence 15 from patent US 6203979.  
AR142617  
AR142617.1 GI:15103903  
Unknown.  
Unknown.  
Unclassified.  
1 (bases 1 to 1081)  
Bandman, O., Hillman, J.L., Yue, H., Guegler, K.J., Corley, N.C.,  
Tang, Y. Tom. and Shah, P.  
Human protease-molecules  
Patent: US 6203979-A 15 20-MAR-2001;

1081 bp  
DNA  
linear  
PAT 08-AUG-2001

FEATURES		Location/Qualifiers	
source		1..1081	
BASE COUNT 202 a 321 c 321 g 237 t			
ORIGIN			
Query Match 100.0%; Score 1081; DB 6; Length 1081;			
Best Local Similarity 100.0%; Pred. No. 5.3e-242;			
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	AGGAGCAGAGGGCGCTCAGGCCGCGGAGAGAGGCCATCGCGCGCGCGGGCGCG 60	
Db	1	AGGAGCAGAGGGCGCTCAGGCCGCGGAGAGAGGCCATCGCGCGCGGGCGCGTG 60	
QY	61	CTGCTGGCGCTCTGCTGGCTCGGCTGACTCAGAGAGCGGAGTTCGAGGAGCGGG 120	
Db	61	CTGCTGGCGCTCTGCTGGCTCGGCTGACTCAGAGAGCGGAGTTCGAGGAGCGGG 120	
QY	121	CCCTTATCAGGACCATGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180	
Db	121	CCCTTATCAGGACCATGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180	
QY	181	CGCGAATCGGGCGCTGGCGCTGGGCTGAGTTCAGAGAGCGGAGTTCGAGGAGCG 240	
Db	181	CGCGAATCGGGCGCTGGCGCTGGGCTGAGTTCAGAGAGCGGAGTTCGAGGAGCG 240	
QY	241	GGAGTGACCTGCTCAGGACCGCTGGGCTGAGTTCAGAGAGCGGAGTTCGAGGAG 300	
Db	241	GGAGTGACCTGCTCAGGACCGCTGGGCTGAGTTCAGAGAGCGGAGTTCGAGGAG 300	
QY	301	AGTGACCTTAGTGATCCCTCGGGTGGAGTTCAGAGTTCGAGGAGTTCGAGGAG 360	
Db	301	AGTGACCTTAGTGATCCCTCGGGTGGAGTTCAGAGTTCGAGGAGTTCGAGGAG 360	
QY	361	TCCTTCTGGAGCGTGCAGGCGCTACTACACCGGTTACTTCGATCGAATATCTACT 420	
Db	361	TCCTTCTGGAGCGTGCAGGCGCTACTACACCGGTTACTTCGATCGAATATCTACT 420	
QY	421	CTCGCTTACCTGGGAAATCACCTATGACATTCCTTGGTGAAGCTGCTGCACCTG 480	
Db	421	CTCGCTTACCTGGGAAATCACCTATGACATTCCTTGGTGAAGCTGCTGCACCTG 480	
QY	481	ACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTGAG 540	
Db	481	ACCTACACTAAACACATCCAGCCCATCTGCTCCAGGCTCCACATTTGAGTTGAG 540	
QY	541	CGACAGACTGCTGGGTGACTGCTGGGGTACATCAAGAGGATGAGGAGTTCGAC 600	
Db	541	CGACAGACTGCTGGGTGACTGCTGGGGTACATCAAGAGGATGAGGAGTTCGAC 600	
QY	601	CCCCACACCTCCAGGAGTTCAGTTCGCGCATCAATAACATCTATGTGCAACACCT 660	
Db	601	CCCCACACCTCCAGGAGTTCAGTTCGCGCATCAATAACATCTATGTGCAACACCT 660	
QY	661	TTCTCTCAAGTACAGTTTCGCGAAGGACATCTTTGGAGAGATGTTTGTCTGGCA 720	
Db	661	TTCTCTCAAGTACAGTTTCGCGAAGGACATCTTTGGAGAGATGTTTGTCTGGCA 720	
QY	721	CAAGCGGGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780	
Db	721	CAAGCGGGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780	
QY	781	GGACTGCTGATCAGATTGGAGTTCGAGTGGGAGTGGGCTGCTGCTGCTGCTGCT 840	
Db	781	GGACTGCTGATCAGATTGGAGTTCGAGTGGGAGTGGGCTGCTGCTGCTGCTGCT 840	
QY	841	CCCGGTGTCTACCAATATCAGCACCACCTTTGAGTGGATCCAGAGCTGATGCC 900	
Db	841	CCCGGTGTCTACCAATATCAGCACCACCTTTGAGTGGATCCAGAGCTGATGCC 900	
QY	901	AGTGGCATGCCAGCCAGCCCTCTCGGCACTACTCTTTTCCCTCTCTCTGCGCT 960	
Db	901	AGTGGCATGTCCAGCCAGCCCTCTCGGCACTACTCTTTTCCCTCTCTCTGCGCT 960	

QY	961	CTCCCACTCTCTGGGGCGGCTCTGAGCCTACTGAGCCCATGAGCCCTGGGGCCACTGCCA 1020	
Db	961	CTCCCACTCTCTGGGGCGGCTCTGAGCCTACTGAGCCCATGAGCCCTGGGGCCACTGCCA 1020	
QY	1021	AGTCAGGCGCTGCT 1080	
Db	1021	AGTCAGGCGCTGCT 1080	
QY	1081	C 1081	
Db	1081	C 1081	
RESULT 2			
AX370460			
LOCUS			
DEFINITION Sequence 7 from Patent WO0198503.			
ACCESSION AX370460			
VERSION AX370460.1 GI:18857514			
KEYWORDS			
SOURCE			
ORGANISM			
human.			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
Xiao,Y.			
Regulation of human eosinophil serine protease 1-like enzyme			
Patent: WO 0198503-A 7 27-DEC-2001;			
Bayer Aktiengesellschaft (DE)			
FEATURES			
Location/Qualifiers			
1..1077			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
BASE COUNT 202 a 320 c 314 g 241 t			
ORIGIN			
Query Match 97.6%; Score 1054.8; DB 6; Length 1077;			
Best Local Similarity 99.8%; Pred. No. 7e-236;			
Matches 1056; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	22	GCCTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 81	
Db	1	GCCTGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60	
QY	82	CGGGCTGGACTCAGGAAGCGGAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 141	
Db	61	CGGGCTGGACTCAGGAAGCGGAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 120	
QY	142	CGAGGGTTCATCAGCTCGGCGCATCTGCTGGTGGAGAGGAGGAGGAGGAGGAG 201	
Db	121	CGAGGGTTCATCAGCTCGGCGCATCTGCTGGTGGAGAGGAGGAGGAGGAGGAG 180	
QY	202	TGGCAGGGGAGGAGCTGCGCTCTGGGATTCGCCAGTATCGGAGTGGAGTGGAG 261	
Db	181	TGGCAGGGGAGGAGCTGCGCTCTGGGATTCGCCAGTATCGGAGTGGAGTGGAG 240	
QY	262	CGCTGGGAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 321	
Db	241	CGCTGGGAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300	
QY	322	GGGTGGATGGTCCAGTTCGGCAGCTGACTTCATTCGATTCGATTCGATTCGAT 381	
Db	301	GGGTGGATGGTCCAGTTCGGCAGCTGACTTCATTCGATTCGATTCGATTCGAT 360	
QY	382	TACTACACCCCTTACTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 441	
Db	361	TACTACACCCCTTACTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 420	
QY	442	CCCTATGACATTCGCTGGTGAAGCTGCTGACCTGCTGACCTGCTGACCTGCTG 501	
Db	421	CCCTATGACATTCGCTGGTGAAGCTGCTGACCTGCTGACCTGCTGACCTGCTG 480	



D	B		841	AGCCACCACCTTTGAGTGGATCCAGAAGACTGATGGCCCCAGAGTGCGCATGTCCCAGGCAGAC	900
Q	Y		922	CCCTCCTGCCACCTACTCTTTTTCCCTTCCTCTGGGCTCTCCCACTCTCTGGGGCGGGTC	981
D	b		901	CCCTCCTGGCGCTACTCTTTTTCCCTTCCTCTGGGCTCTCCCACTCTCTGGGGCGGGTC	960
Q	Y		982	TGAGCTTACTGTAGCCCCATCAGCCTGGGGCCACTGCCAAGTCAGGCGCTGTCTCTCTTC	1041
D	b		961	TGAGCTTACTGTAGCCCCATCAGCCTGGGGCCACTGCCAAGTCAGGCGCTGTCTCTCTTC	1020
Q	Y		1042	TGCTCTCTTTGGTAATAAACACATTCAGTTGATGCCT	1079
D	b		1021	TGCTCTCTTTGGTAATAAACACATTCAGTTGATGCCT	1058
 RESULT 4 AX370459					
L	O	CUS	AX370459	Sequence 6 from Patent WO0198503.	
A	C	ESSION	AX370459		
V	E	SION	AX370459.1	GI:18857513	
K	E	WORDS			
S	O	RCE		human.	
O	R	GANISM		Homo sapiens	
R	E	FERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
A	U	TORS		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
T	I	TLE		xiao, Y.	
J	O	URNAL		Regulation of human eosinophil serine protease 1-like enzyme	
F	E	URES		Patent: WO 0198503-A 6 27-DEC-2001;	
	s	rce		Bayer Aktiengesellschaft (DE)	
				Location/Qualifiers	
				1..1082	
				/organism="Homo sapiens"	
				/db_xref="taxon:9606"	
B	A	S	COUNT	216 a 318 c 308 g 240 t	
O	R	I	GIN		
 Query Match                  96.7%; Score 1045.8; DB 6; Length 1082; Best Local Similarity      99.8%; Pred. No. 8.9e-234; Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
Q	Y		31	GAGGAGGCCATGGCGCGCGCGGGCGGTGCTGTGTGGCGTGTGTGTGGCTCGGGCTGGA	90
D	b		1	GAGGAGGCCATGGCGCGCGCGGGCGGTGCTGTGTGGCGTGTGTGTGGCTCGGGCTGGA	60
Q	Y		91	CTCAGGAAGCCGGAGTGTGCAGGAGCGCGCCCTTATCAGGACCATCGCGCCAGCGGTC	150
D	b		61	CTCAGGAAGCCGGAGTGTGCAGGAGCGCGCCCTTATCAGGACCATCGCGCCAGCGGTC	120
Q	Y		151	ATCAGTTCGCCCATCTGTGGTGGAGAGACCGCGAACTCGGGCGTTGGCGGTGGCACGGG	210
D	b		121	ATCAGTTCGCCCATCTGTGGTGGAGAGAGACCGCGAACTCGGGCGTTGGCGGTGGCACGGG	180
Q	Y		211	AGCTTGGCGCTTGGGAATCCCACGATGCGGAGTAGGCGTGTCTAGCCACCGCTGGGCA	270
D	b		181	AGCTTGGCGCTTGGGAATCCCACGATGCGGAGTAGGCGTGTCTAGCCACCGCTGGGCA	240
Q	Y		271	CTCAGCGCGCGCACTGCTTTGAACTATAGTAGACCTTAGTAGATCCCTCCGGGTGGATG	330
D	b		241	CTCAGCGCGCGCACTGCTTTGAACTATAGTAGACCTTAGTAGATCCCTCCGGGTGGATG	300
Q	Y		331	GTCAGTTTGGCAGCTGTACTTCCATGCCATCTCTTGGAGCGCTCGAGCGCTTACTACAC	390
D	b		301	GTCAGTTTGGCAGCTGTACTTCCATGCCATCTCTTGGAGCGCTCGAGCGCTTACTACAC	360
Q	Y		391	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCAACCTATGAC	450
D	b		361	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCAACCTATGAC	420
Q	Y		451	ATTGCGCTTGGTGAAGCTGTCTGCACTGTGCACCTTACACTTAACACATCCAGCCATCTGT	510
D	b		421	ATTGCGCTTGGTGAAGCTGTCTGCACTGTGCACCTTACACTTAACACATCCAGCCATCTGC	480

/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 216 a 318 c 308 g 240 t  
ORIGIN  
Query Match 96.7%; Score 1045.8; DB 6; Length 1082;  
Best Local Similarity 99.8%; Pred. No. 8.9e-234;  
Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 31 GAGAGGCCATGGCGCGCGCGCGCGCTGCTGGCGCTGCTGGCGCTGCTGGCGCTGGA 90  
Db 1 GAGAGGCCATGGCGCGCGCGCGCGCTGCTGGCGCTGCTGGCGCTGCTGGCGCTGGA 60  
QY 91 CTCAGGAAGCGGAGTCCAGAGCGCGCGCTTATCAGGACCATGGCGCGCGGCTC 150  
Db 61 CTCAGGAAGCGGAGTCCAGAGCGCGCGCTTATCAGGACCATGGCGCGCGGCTC 120  
QY 151 ATCAGCTCGGCATCGTGGTGGAGAGACGCCGAACCTCGGCGCTGGCCGTGGCAGGG 210  
Db 121 ATCAGCTCGGCATCGTGGTGGAGAGACGCCGAACCTCGGCGCTGGCCGTGGCAGGG 180  
QY 211 AGCCTGGCGCTGGGATCCACGATGCGGAGTGGAGTGGCTCAGCCACCGCTGGGCA 270  
Db 181 AGCCTGGCGCTGGGATCCACGATGCGGAGTGGAGTGGCTCAGCCACCGCTGGGCA 240  
QY 271 CTCAGGCGCGCGCTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCGGCGTGGATG 330  
Db 241 CTCAGGCGCGCGCTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCGGCGTGGATG 300  
QY 331 GTCCAGTTGGCCAGCTGACATCCATGCCATCTTCCTGGAGCCCTGCAGGCGCTACTACACC 390  
Db 301 GTCCAGTTGGCCAGCTGACATCCATGCCATCTTCCTGGAGCCCTGCAGGCGCTACTACACC 360  
QY 391 CGTTACTCTGATCGAATATCTAGTACCTGTACCTTACCTAAACACATCCAGCCCATCTGT 450  
Db 361 CGTTACTCTGATCGAATATCTAGTACCTGTACCTTACCTAAACACATCCAGCCCATCTGC 420  
QY 451 ATTCCCTTGGTGAAGCTGTCTGCACTGTACCTTACCTAAACACATCCAGCCCATCTGT 510  
Db 421 ATTCCCTTGGTGAAGCTGTCTGCACTGTACCTTACCTAAACACATCCAGCCCATCTGC 480  
QY 511 CTCAGGCGCTCCAGATTTGAGAACCGGACAGACTGCTGGGTGACATGGGTGGGG 570  
Db 481 CTCAGGCGCTCCAGATTTGAGAACCGGACAGACTGCTGGGTGACATGGGTGGGG 540  
QY 571 TACATCAAGAGGATGAGGCACTGCCATCTCCACACCCCTCCAGGAGTTTCAGGTCGCC 630  
Db 541 TACATCAAGAGGATGAGGCACTGCCATCTCCACACCCCTCCAGGAGTTTCAGGTCGCC 600  
QY 631 ATCATAAACCACTATGTGCAACACACCTCTTCCTCAAGTACAGTTTCGCAAGGACATC 690  
Db 601 ATCATAAACCACTATGTGCAACACACCTCTTCCTCAAGTACAGTTTCGCAAGGACATC 660  
QY 691 TTTGAGACATGTTGTGCTGGCAATGCCCAAGCGGGAGAGATGCTGCTTCGGTGAC 750  
Db 661 TTTGAGACATGTTGTGCTGGCAATGCCCAAGCGGGAGAGATGCTGCTTCGGTGAC 720  
QY 751 TCAGGTGACCTTTGGCTGTAAAGAAATGGACTGTGGTATCAGATTGGAGTCTGTAGC 810  
Db 721 TCAGGTGACCTTTGGCTGTAAAGAAATGGACTGTGGTATCAGATTGGAGTCTGTAGC 780  
QY 811 TGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGCTGTCTACCAATATCAGCCACAC 870  
Db 781 TGGGAGTGGGCTGTGGTGGCCCAATCGGCCCGCTGTCTACCAATATCAGCCACAC 840  
QY 871 TTTGAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGTCCAGCCAGACCCCTCCCTGG 930  
Db 841 TTTGAGTGGATCCAGAGCTGATGGCCAGAGTGGCATGTCCAGCCAGACCCCTCCCTGG 900  
QY 931 CCAGTACTCTTTTCCCTTCTCTGGGCTCTCCACTCTCCAGGCGCGGCTGTGAGCCTAC 990  
Db 901 CCAGTACTCTTTTCCCTTCTCTGGGCTCTCCACTCTCCAGGCGCGGCTGTGAGCCTAC 960

QY 991 CTGAGCCCATGAGCGCTGGCGCGCGCTGCAAGTCAAGCCCTGGTCTCTCTCTGCTTGT 1050  
Db 961 CTGAGCCCATGAGCGCTGGCGCGCGCTGCAAGTCAAGCCCTGGTCTCTCTCTGCTTGT 1020  
QY 1051 TGGTAATAAACACATTCACAGTTGATGCT 1079  
Db 1021 TGGTAATAAACACATTCACAGTTGATGCT 1049  
RESULT 6  
E21867  
LOCUS E21867 1082 bp DNA linear PAT 18-JUN-2001  
DEFINITION Novel acidophil serine protease.  
ACCESSION E21867  
VERSION E21867.1 GI:13023738  
KEYWORDS JP 1999032768-A/3.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1082)  
AUTHORS Hiroshi, K. and Masahiro, I.  
TITLE Novel acidophil serine protease  
JOURNAL Patent: JP 1999032768-A 3 09-FEB-1999;  
ONO PHARMACEUT CO LTD  
COMMENT OS Homo Sapiens  
PN JP 1999032768-A/3  
PD 09-FEB-1999  
PF 16-JUL-1997 JP 1997191319  
PR HIROSHI KIDO, MASAHIRO INOUE  
PC C12N15/09, A61K38/55, A61K39/395, A61K48/00, C07K7/00,  
C07K16/40  
PC C12N9/64, C12N15/00, A61K37/64, A61K37/64  
CC Topology: Linear;  
FH Key Location/Qualifiers  
FT sig peptide 10..951  
FT mat peptide 133..951.  
FEATURES  
source  
1..1082  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 216 a 318 c 308 g 240 t  
ORIGIN  
Query Match 96.7%; Score 1045.8; DB 6; Length 1082;  
Best Local Similarity 99.8%; Pred. No. 8.9e-234;  
Matches 1047; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 31 GAGAGGCCATGGCGCGCGCGCGCTGCTGGCGCTGCTGGCGCTGCTGGCGCTGGA 90  
Db 1 GAGAGGCCATGGCGCGCGCGCGCGCTGCTGGCGCTGCTGGCGCTGCTGGCGCTGGA 60  
QY 91 CTCAGGAAGCGGAGTCCAGAGCGCGCGCTTATCAGGACCATGGCGCGCGGCTC 150  
Db 61 CTCAGGAAGCGGAGTCCAGAGCGCGCGCTTATCAGGACCATGGCGCGCGGCTC 120  
QY 151 ATCAGCTCGGCATCGTGGTGGAGAGACGCCGAACCTCGGCGCTGGCCGTGGCAGGG 210  
Db 121 ATCAGCTCGGCATCGTGGTGGAGAGACGCCGAACCTCGGCGCTGGCCGTGGCAGGG 180  
QY 211 AGCCTGGCGCTGGGATCCACGATGCGGAGTGGAGTGGCTCAGCCACCGCTGGGCA 270  
Db 181 AGCCTGGCGCTGGGATCCACGATGCGGAGTGGAGTGGCTCAGCCACCGCTGGGCA 240  
QY 271 CTCAGGCGCGCGCTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCGGCGTGGATG 330  
Db 241 CTCAGGCGCGCGCTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCGGCGTGGATG 300  
QY 331 GTCCAGTTGGCCAGCTGACATCCATGCCATCTTCCTGGAGCCCTGCAGGCGCTACTACACC 390  
Db 301 GTCCAGTTGGCCAGCTGACATCCATGCCATCTTCCTGGAGCCCTGCAGGCGCTACTACACC 360

QY	391	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCATGAC	450
Db	361	CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCATGAC	420
QY	451	ATTGCCCTTGGTGAAGCTGTCTGCACCTGTCACTTAAACACATTCAGGCCATCTGT	510
Db	421	ATTGCCCTTGGTGAAGCTGTCTGCACCTGTCACTTAAACACATTCAGGCCATCTGC	480
QY	511	CTCCAGGCCCTCCACATTTGACTTTGAGAACGGACAGACTGCTGGGTGACTGCGCTGGGG	570
Db	481	CTCCAGGCCCTCCACATTTGACTTTGAGAACGGACAGACTGCTGGGTGACTGCGCTGGGG	540
QY	571	TACATCAAGAGGATGAGGCACTGCCATCTCCCCACACCTCCAGAAAGTTTCAAGTTCGCC	630
Db	541	TACATCAAGAGGATGAGGCACTGCCATCTCCCCACACCTCCAGAAAGTTTCAAGTTCGCC	600
QY	631	ATCATAAACAACCTATGTGCAACACCTCTTCTCAAGTACAGTTTCCGCAAGGACATC	690
Db	601	ATCATAAACAACCTATGTGCAACACCTCTTCTCAAGTACAGTTTCCGCAAGGACATC	660
QY	691	TTTGGAGACATGGTTTGTGCTGGCAATGCCAAGCGGGAAGGATGCCTTCCGTCGAGC	750
Db	661	TTTGGAGACATGGTTTGTGCTGGCAATGCCAAGCGGGAAGGATGCCTTCCGTCGAGC	720
QY	751	TCAGTGGACCTTTGGCCTGTAAACAAGATGGACTGTGTATCAGATTTGGAGTGTGAGC	810
Db	721	TCAGTGGACCTTTGGCCTGTAAACAAGATGGACTGTGTATCAGATTTGGAGTGTGAGC	780
QY	811	TGGGAGTGGGCTGTGTCGGCCCAATCGGCCGGTGTCTACACCAATATCAGCCACCAC	870
Db	781	TGGGAGTGGGCTGTGTCGGCCCAATCGGCCGGTGTCTACACCAATATCAGCCACCAC	840
QY	871	TTTCAGTGGATCCAGAAGCTGATGGCCCAAGATGGCATGTCCCAGCAGACCCCTCTGG	930
Db	841	TTTCAGTGGATCCAGAAGCTGATGGCCCAAGATGGCATGTCCCAGCAGACCCCTCTGG	900
QY	931	CCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGGCGGGTGTGAGCCCTAC	990
Db	901	CCACTACTCTTTTCCCTCTCTCTGGGCTCTCCACTCTCTGGGGCGGGTGTGAGCCCTAC	960
QY	991	CTGAGCCCATGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCCTCTCTCTGTGT	1050
Db	961	CTGAGCCCATGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCCTCTCTCTGTGT	1020
QY	1051	TGCTAATAACACATTCAGTTGATGCCT	1079
Db	1021	TGCTAATAACACATTCAGTTGATGCCT	1049

RESULT 7	
AB031329	
LOCUS	1082 bp mRNA linear PRI 15-JAN-2000
DEFINITION	Homo sapiens esp-1 mRNA for eosinophil serine protease, complete cds.
ACCESSION	AB031329
VERSION	AB031329.1 GI:5777329
KEYWORDS	eosinophil serine protease.
SOURCE	Homo sapiens male eosinophil cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (sites)
TITLE	Inoue, M., Kanbe, N., Kurosawa, M. and Kido, H.
	Cloning and tissue distribution of a novel serine protease esp-1 from human eosinophils
JOURNAL	Biochem. Biophys. Res. Commun. 252 (2), 307-312 (1998)
MEDLINE	9904501
REFERENCE	2 (sites)
AUTHORS	Inoue, M., Isobe, M., Itoyama, T. and Kido, H.
TITLE	Structural analysis of esp-1 gene (PRSS 21)
JOURNAL	Biochem. Biophys. Res. Commun. 266 (2), 564-568 (1999)
MEDLINE	2006805

REFERENCE 3 (bases 1 to 1082)  
AUTHORS Inoue,M., Kido,H., Kanbe,N. and Kurosawa,M.  
TITLE Direct Submission  
JOURNAL Submitted (20-AUG-1999) Masahiro Inoue, Institute for Enzyme Research, Division of Enzyme Chemistry; 3 Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan (E-mail:inoue@ier.tokushima-u.ac.jp, Tel:81-886-33-7424, Fax:81-886-33-7425)

FEATURES  
source Location/Qualifiers  
1..1082  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="16p13.3"  
/chromosome="16"  
/sex="male"  
/tissue\_type="eosinophil"  
gene 1..1082  
/gene="esp-1"  
CDS 10..954  
/gene="esp-1"  
/codon\_start=1  
/product="eosinophil serine protease"  
/protein\_id="BAAB3520.1"  
/db\_xref="GI:5777330"  
/translation="MGARGALLLALLIARAGLRKPSOEAAPLSGPCGRVITSRIVG  
GEAELRWPFQWSRLWDSHGVSLLSHRWALTAAHFETVSDLSDPGMVQFGQ  
LTPSMFSWLQAVYTRFVSNLYSPRYLGNSPYDIALVKLSAPVTTHKIQCICQA  
STFEFNRTDCWTGWGYIKEDEALSPHTLQEVQVAIIINSMCNHLFLKYISFRKDIF  
GDWVCAGNAOGGKADCFDGSGGPLNKNGLWTOIGVSWGVCGCRPNRPVYTNIISH  
HFWEIKMLAQSGMSQPDPSWPLLFPELLMALPLLGV"  
polyA\_site 1064  
/gene="esp-1"  
/note="18 a nucleotides"  
BASE COUNT 216 a 318 c 308 g 240 t  
ORIGIN

Query Match	96.7%	Score 1045.8;	DB 9;	Length 1082;
Best Local Similarity	99.8%;	Pred. No. 8.9e-234;		
Matches 1047;	Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;

```

QY   31  GAGGAGGCCATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGCTGGCTGGA  90
Db    1  GAGGAGGCCATGGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTGGCTGGA  60

QY   91  CTCAGGAAGCCGGAGTCGCAGGAGCGGGCGCCCTTATCAGGACCATCGCGCCGACGGGTC  150
Db    61  CTCAGGAAGCCGGAGTCGCAGGAGCGGGCGCCCTTATCAGGACCATCGCGCCGACGGGTC  120

QY   151  ATCACGTCCGCATCTGGGTGGAGAGAGCGCGAACTCGGGCGTTGGCCGTGGCAGGG  210
Db    121  ATCACGTCCGCATCTGGGTGGAGAGAGCGCGAACTCGGGCGTTGGCCGTGGCAGGG  180

QY   211  AGCCTCGGCTGTGGGATTCCCACGATATCCGAGTGAGCGAGTGCAGCCACCGCTGGGCA  270
Db    181  AGCCTCGGCTGTGGGATTCCCACGATATCCGAGTGAGCGAGTGCAGCCACCGCTGGGCA  240

QY   271  CTCACGGCGGCGACTGCTTTGAAAACCTATAGTAGACCTTAGTGATCCCTCCGGGTGGATG  330
Db    241  CTCACGGCGGCGACTGCTTTGAAAACCTATAGTAGACCTTAGTGATCCCTCCGGGTGGATG  300

QY   331  GTCCAGTTGGCCAGCTGACTTCCATGCCATCTTCTTGAGCGCTGCAGGCCCTACTACACC  390
Db    301  GTCCAGTTGGCCAGCTGACTTCCATGCCATCTTCTTGAGCGCTGCAGGCCCTACTACACC  360

QY   391  CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTACCCCTATGAC  450
Db    361  CGTTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTACCCCTATGAC  420

QY   451  ATTGCTTGGTGAAGCTGCTGCACCTGTACCTACACTAACACATCCAGCCCATCTGT  510
Db    421  ATTGCTTGGTGAAGCTGCTGCACCTGTACCTACACTAACACATCCAGCCCATCTGC  480

QY   511  CTCACGGCCTCCACATTTGAGTTTGAGAACGGACAGACTGCTGGGTGACTGGCTGGGG  570

```





```

Db 721 GGACCTTGGCTGTAAACAAGATGGAGCTGTGGTATCAGATTGGAGTCTGTAGCTGGGA 780
QY 817 GTGGGCTGTGGCGCCCAATCGGCCGCTGTCTACACCAATATCAGCCACCACTTTGAG 876
Db 781 GTGGGCTGTGGCGCCCAATCGGCCGCTGTCTACACCAATATCAGCCACCACTTTGAG 840
QY 877 TGGATCCAGAAGCTGATGGCCGAGAGTGGCAGTGTCCAGCCAGACCCCTCTGGCCACTA 936
Db 841 TGGATCCAGAAGCTGATGGCCGAGAGTGGCAGTGTCCAGCCAGACCCCTCTGGCCGCTA 900
QY 937 CTCCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCCCACTCTGGGGCGGCTGAGCCTGAGC 996
Db 901 CTCCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCCCACTCTGGGGCGGCTGAGCCTGAGC 960
QY 997 CCATGACGCTTGGGCCACTGCCAAGTCCAGGCTGGTCTCTCTCTCTCTCTCTCTCTCTCT 1056
Db 961 CCATGACGCTTGGGCCACTGCCAAGTCCAGGCTGGTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1057 TAAACACATTCAGTTGATGCT 1079
Db 1021 TAAACACATTCAGTTGATGCT 1043

RESULT 9
E21853
LOCUS E21853 1085 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel acidophil serine protease.
ACCESSION E21853
VERSION E21853.1 GI:13023724
KEYWORDS JP 1999032768-A/5.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1085)
AUTHORS Hiroshi,K. and Masahiro,I.
TITLE Novel acidophil serine protease
JOURNAL Patent: JP 1999032768-A 5 09-FEB-1999;
ONOPHARMACEUT CO LTD
COMMENT
PS Unidentified
PN JP 1999032768-A/5
PD 09-FEB-1999
PF 16-JUL-1997 JP 1997191319
PR
PI HIROSHI KIDO, MASAHIRO INOUE
PC C12N15/09,A61K38/55,A61K39/395,A61K48/00,C07K1/00,
PC C07K16/40,
PC C12N9/64,C12N15/00,A61K37/64,A61K37/64
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source
FT 1..1085
FT Location/Qualifiers
FT 1..1085
FT /organism='Unidentified'.
FT /organism='unidentified'
FT /db_xref='taxon:32644'
BASE COUNT 217 a 318 c 308 g 242 t
ORIGIN

Query Match 95.7%; Score 1034.4; DB 6; Length 1085;
Best Local Similarity 99.6%; Pred. No. 4.1e-231;
Matches 1048; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 31 GAGGAGCCATGGGCGCGGCGCTGCTGTGGCGCTGCTGTGGCGCTGCTGTGGCGCTGGA 90
Db 1 GAGGAGCCATGGGCGCGGCGGCGCTGCTGTGGCGCTGCTGTGGCGCTGGA 60
QY 91 CTCAGGAAGCCGAGTCCAGAGGCGCGCCCTATCAGACCATGCGCGCGAGGCTC 150
Db 61 CTCAGGAAGCCGAGTCCAGAGGCGCGCGCTTATCAGACCATGCGCGCGAGGCTC 120
QY 151 ATCAGCTGGCGATCGTGGTGGAGAGGACCGCAACTCGGGCGCTGGCGCTGGCAGGG 210

```

```

Db 121 ATCAGCTCCGCGCATCGTGGTGGAGAGGAGCGCGAATCTGGGCGTGTGGCGGTGGCAGGG 180
QY 211 AGCCTGCGCCTGTGGGATTCACACGATATCGGAGTGAGCCTGCTCAGCCACCGCTGGGCA 270
Db 181 AGCCTGCGCCTGTGGGATTCACACGATATCGGAGTGAGCCTGCTCAGCCACCGCTGGGCA 240
QY 271 CTCAGGGGGGCGCACTGCTTTGAAACCTATAGTAGACCTTAGTATCCCTCCGGGTGGATG 330
Db 241 CTCAGGGGGGCGCACTGCTTTGAAACCTATAGTAGACCTTAGTATCCCTCCGGGTGGATG 300
QY 331 GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCCTTCTGGAGCCTGAGGCTTACTACACC 390
Db 301 GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCCTTCTGGAGCCTGAGGCTTACTACACC 360
QY 391 CGTTACTTCGATCGAATATATCTATCTGAGCCTCTGCTACCTGGGAAATTCACCTTATCAC 450
Db 361 CGTTACTTCGATCGAATATATCTATCTGAGCCTCTGCTACCTGGGAAATTCACCTTATCAC 420
QY 451 ATTGCTTGGTGAAGCTGTCTGCACTGTACCTAC---ACTAAACATCCAGCCCATC 507
Db 421 ATTGCTTGGTGAAGCTGTCTGCACTGTACCTACCTACCTAAACATCCAGCCCATC 480
QY 508 TGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGG 567
Db 481 TGTCTCCAGGCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGG 540
QY 568 GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACCCCTCCAGGAAGTTTCAGGTC 627
Db 541 GGGTACATCAAGAGGATGAGGCACTGCCATCTCCCAACCCCTCCAGGAAGTTTCAGGTC 600
QY 628 GCCATCAAAACACTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCGAAGGAC 687
Db 601 GCCATCAAAACACTCTATGTGCAACCACTCTTCTCTCAAGTACAGTTTCGCGAAGGAC 660
QY 688 ATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGCGGAGGATGCTGCTTCGGT 747
Db 661 ATCTTTGGAGACATGTTTGTGCTGCAATGCCAAGCGGAGGATGCTGCTTCGGT 720
QY 748 GACTAGTGGACCCCTTGCCCTGTAAAGAATGAGTGTGTATCAGATTGGAGTCGTG 807
Db 721 GACTAGTGGACCCCTTGCCCTGTAAAGAATGAGTGTGTATCAGATTGGAGTCGTG 780
QY 808 AGCTGGGAGTGGCTGTGTCGCCCAATCGGCCGCTGTCTACACCAATATCAGCCAC 867
Db 781 AGCTGGGAGTGGCTGTGTCGCCCAATCGGCCGCTGTCTACACCAATATCAGCCAC 840
QY 868 CACTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGTCCAGCAGACCCCTCC 927
Db 841 CACTTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGTCCAGCAGACCCCTCC 900
QY 928 TGGCCACTACTCTTTTCCCTCTCTGGGCTCTCCCACTCTGGGCGGCTGTGAGCC 987
Db 901 TGGCCACTACTCTTTTCCCTCTCTGGGCTCTCCCACTCTGGGCGGCTGTGAGCC 960
QY 988 TACCTGAGCCCATGAGCTGGGCGCACTGCAAGTCCAGGCTGCTCTCTCTCTCTCTCTCT 1047
Db 961 TACCTGAGCCCATGAGCTGGGCGCACTGCAAGTCCAGGCTGCTCTCTCTCTCTCTCTCT 1020
QY 1048 GTTGGTGAATAAACACATTCAGTTGATGCT 1079
Db 1021 GTTGGTGAATAAACACATTCAGTTGATGCT 1052

RESULT 10
E21854
LOCUS E21854 1085 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel acidophil serine protease.
ACCESSION E21854
VERSION E21854.1 GI:13023725
KEYWORDS JP 1999032768-A/6.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Db	601	GCATCATAAACAACACTTATGTGCAACCACTCTCTCTCAAGTACAGTTTCCGCAAGGAC	660
QY	688	ATCTTTGGAGACATGGTTTGTCTGGCAATGCCAAAGCGGGAAGGATGCTGTCTCGGT	747
Db	661	ATCTTTGGAGACATGGTTTGTCTGGCAATGCCAAAGCGGGAAGGATGCTGTCTCGGT	720
QY	748	GACTCAGGTGGACCCCTTGGCCCTGTAACAAGAAATGGACTGTGTATCAGATGGAGTCGTG	807
Db	721	GACTCAGGTGGACCCCTTGGCCCTGTAACAAGAAATGGACTGTGTATCAGATGGAGTCGTG	780
QY	808	AGCTGGGAGTGGCTGTGGTGCSCCCCAATCGGCCCGGTGTCTACACAAATATCAGCCAC	867
Db	781	AGCTGGGAGTGGCTGTGGTGCSCCCCAATCGGCCCGGTGTCTACACAAATATCAGCCAC	840
QY	868	CACCTTTGAGTGGATCCAGAAGCTGATGCCCGAGAGTGGCATGTCCAGCCAGACCCCTCC	927
Db	841	CACCTTTGAGTGGATCCAGAAGCTGATGCCCGAGAGTGGCATGTCCAGCCAGACCCCTCC	900
QY	928	TGGCCACTACTCTTTTTCCTCTCTCTGCGGCTCTCCCACTCTCTGGGCGCGGTCTGAGCC	987
Db	901	TGGCCACTACTCTTTTTCCTCTCTCTGCGGCTCTCCCACTCTCTGGGCGCGGTCTGAGCC	960
QY	988	TACCTGAGCCCATCGAGCTGGGCGCACTGCCAAGTCAAGCTCAGGCGCTCTCTCTCTCTT	1047
Db	961	TACCTGAGCCCATCGAGCTGGGCGCACTGCCAAGTCAAGCTCAGGCGCTCTCTCTCTT	1020
QY	1048	GTTTGGTAATAAACACATTCAGTTGATGCCT	1079
Db	1021	GTTTGGTAATAAACACATTCAGTTGATGCCT	1052
RESULT 11			
LOCUS	AB031331		
DEFINITION	Homo sapiens tesp-1 mRNA for eosinophil serine protease 1 splicing variant, complete cds.	1119 bp	mRNA linear PRI 18-JAN-2000
ACCESSION	AB031331		
VERSION	AB031331.1	GI:6714621	
KEYWORDS	tesp-1; eosinophil serine protease 1 splicing variant.		
SOURCE	Homo sapiens	cell_line:HeLa S3	CDNA to mRNA, clone_lib:Human HeLa cDNA.
ORGANISM	Homo sapiens		
REFERENCE	1 (sites)		
AUTHORS	Inoue, M., Isobe, M., Itoyama, T. and Kido, H.		
TITLE	Structural analysis of esp-1 gene (PRSS 21)		
JOURNAL	Biochem. Biophys. Res. Commun. 266 (2), 564-568 (1999)		
MEDLINE	20069805		
REFERENCE	2 (bases 1 to 1119)		
AUTHORS	Inoue, M., Kido, H. and Isobe, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-AUG-1999) Masahiro Inoue, Institute for Enzyme Research, Division of Enzyme Chemistry; 3 Kuramoto-cho, Tokushima, Tokushima 770-8503, Japan (E-mail: inoue@ier.tokushima-u.ac.jp, Tel: 81-886-33-7424, Fax: 81-886-33-7425)		
FEATURES	Location/Qualifiers		
source	1..1119		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="16"		
	/map="16p13.3"		
	/cell_line="HeLa S3"		
	/clone_lib="Human HeLa cDNA"		
gene	1..1119		
	/gene="tesp-1"		
CDS	90..992		
	/gene="tesp-1"		
	/codon_start=1		
	/product="eosinophil serine protease 1 splicing variant"		
	/protein_id="BAA89532.1"		
	/db_xref="GI:6714622"		
	/translation="MGARGALLALLIARAGLRKPEQAEAPLGGCGRRVTVSRVIGV"		

BASE COUNT  
ORIGIN

Query Match	89.7%;	Score 969.2;	DB 9;	Length 1119;
Best Local Similarity	95.7%;	Pred. No. 7e-216;		
Matches 1034; Conservative	0;	Mismatches 3;	Indels 43;	Cases 2;

QY	1	AGAGCAGAGGGGGCGT	CAGCCCGGGAGAGGAGCCATGGGCGCGCGGGGCGGTG	60
Db				
Db	51	AGGAGCAGAGGGGGCGT	CAGCCCGGGAGAGGAGCCATGGGCGCGGGGCGGTG	110
QY	61	CTGCTGGCGCTGCTGCT	GGCGCTGAGACTCAGAGCCGGAGTCCGAGGAGCGGCG	120
Db				
Db	111	CTGCTGGCGCTGCTGCT	GGCGCTGAGACTCAGAGCCGGAGTCCGAGGAGCGGCG	170
QY	121	CCCTTATCAGGACCATTG	CGCGCGAGGGTTCATCAGTCCGCATCGTGGGTGGAGAGAC	180
Db				
Db	171	CCGTTATCAGGACCATTG	CGCGCGAGGGTTCATCAGTCCGCATCGTGGGTGGAGAGAC	230
QY	181	GGCGAACTCGGGCGTTG	GGCCGTGGCAGGGAGCCCTGCGCTGTGGATTCGCCACGTATGC	240
Db				
Db	231	GGCGAACTCGGGCGTTG	GGCCGTGGCAGGGAGCCCTGCGCTGTGGATTCGCCACGTATGC	290
QY	241	GGAGTGAGCGCTCAGCA	CCCGTGGGCACCTCAGCGGGCGCACTGCTTTGAAACCTAT	300
Db				
Db	291	GGAGTGAGCGCTCAGCA	CCCGTGGGCACCTCAGCGGGCGCACTGCTTTGAAACCTAT	350
QY	301	AGTGACCTTAGTGATCC	TCCGGGTGGATGGTCAGATTTGGCCAGCTGACTTCATGCCA	360
Db				
Db	351	AGTGACCTTAGTGATCC	TCCGGGTGGATGGTCAGATTTGGCCAGCTGACTTCATGCCA	410
QY	361	TCCCTCTGGAGCCCTG	CAGCGCTACTACACCGTTACTTCGTAATCTATCTGAGC	420
Db				
Db	411	TCCCTCTGGAGCCCTG	CAGCGCTACTACACCGTTACTTCGTAATCTATCTGAGC	470
QY	421	CCTCGCTACCTTGGGA	ATTACCCCTATGACATTTGCCCTTGGTGAAGCTGTGCACTGTC	480
Db				
Db	471	CCTCGCTACCTTGGGA	ATTACCCCTATGACATTTGCCCTTGGTGAAGCTGTGCACTGTC	530
QY	481	ACCTACACTAAACACAT	CCAGCCCATCTGCTCCAGGCGCTCCACATTTGATTTGAGAAC	540
Db				
Db	531	ACCTACACTAAACACAT	CCAGCCCATCTGCTCCAGGCGCTCCACATTTGATTTGAGAAC	590
QY	541	CGGACAGACTGCTGGGT	GACTGGCTGGGGTACATCAAGAGGATGAGGCACTGCCATCT	600
Db				
Db	591	CGGACAGACTGCTGGGT	GACTGGCTGGGGTACATCAAGAGGATGAGGCACTGCCATCT	650
QY	601	CCCCACACCTCCAGGA	AGTTCCAGTCCGATCAATCAACAACTCTATGTCACACCACTC	660
Db				
Db	651	CCCCACACCTCCAGGA	AGTTCCAGTCCGATCAATCAACAACTCTATGTCACACCACTC	710
QY	661	TTCTCTAAGTACAGTTT	CCCAAGGACATCTTTGGAGACATGTTTGTGCTGGCAATGCC	720
Db				
Db	711	TTCTCTAAGTACAGTTT	CCCAAGGACATCTTTGGAGACATGTTTGTGCTGGCAATGCC	770
QY	721	CAAGGGGGGAAGGATCG	CTGCTTCGGTGACTCAGGTGGACCCCTTGCCCTGTAAACAAGAT	780
Db				
Db	751	-----GGTGACTCAGGT	GGACCCCTTGCCCTGTAAACAAGAT	788
QY	781	GGAGTGTGGTATCAGAT	TGGAGTCGTGAGCTGGGGAGTGGCTGTGGTGGGCCCAATCGG	840
Db				
Db	789	GGAGTGTGGTATCAGAT	TGGAGTCGTGAGCTGGGGAGTGGCTGTGGTGGGCCCAATCGG	848
QY	841	CCGGGTGTCTACACCA	ATATCAGCCACCACTTTGAGTGGATCCAGAAGCTGATGCCACG	900
Db				
Db	849	CCGGGTGTCTACACCA	ATATCAGCCACCACTTTGAGTGGATCCAGAAGCTGATGCCACG	908
QY	901	AGTGGGATGTCCAGCA	CGACCCCTCTCTGGGCACCTACTCTTTTCCCTCTCTCTCTGGGCT	960

Db	909	AGTGGGATGCCAGCAGACCCCTCTCGCGGTACTCTTTTCCTCTTGCGCT	968
QY	961	CTCCCCACTCTGGGGCGGTCTGAGCTACCTAGCCCATGAGCCTGGGGCACTGCCA	1020
Db	969	CTCCCACTCTGGGGCGGTCTGAGCCTACCTAGCCCATGAGCCTGGGGCACTGCCA	1028
QY	1021	AGTCAGGCCTGGTTCT-CTTCTGTCTGTTTTGGTAATAAACACATTCCAGTTGATGCT	1079
Db	1029	AGTCAGGCCTGGTTCTACTTCTGTCTGTTGGTAAATAACACATTCCAGTTGATGCT	1088
 RESULT 12			
E21865			
LOCUS			
DEFINITION Novel acidophil serine protease.			
ACCESSION E21865			
VERSION E21865.1 GI:13023736			
KEYWORDS JP 1999032768-A/1.			
SOURCE unclassified.			
ORGANISM unclassified.			
REFERENCE			
AUTHORS 1 (bases 1 to 942)			
TITLE Hiroshi,K. and Masahiro,I.			
JOURNAL Novel acidophil serine protease			
COMMENT Patent: JP 1999032768-A 1 09-FEB-1999;			
ONO PHARMACEUT CO LTD			
OS Unidentified			
PN JP 1999032768-A/1			
PD 09-FEB-1999			
PR 16-JUL-1997 JP 1997191319			
PI HIROSHI KIDO,MASAHIRO INOUE			
PC C12N15/09,A61K38/55,A61K38/55,A61K39/395,A61K48/00,C07K7/00,			
PC C07K16/40			
PC C12N9/64,C12N15/00,A61K37/64,A61K37/64			
CC Strandedness: Single;			
CC Topology: Linear;			
FH Key Location/Qualifiers			
FT source 1..942			
FT Location/Qualifiers			
source .1..942			
/organism='Unidentified'.			
/db_xref="taxon:32644"			
BASE COUNT 176 a 283 c 276 g 207 t			
ORIGIN			
 Query Match 86.8%; Score 938.8; DB 6; Length 942;			
Best Local Similarity 99.8%; Pred. No. 8.9e-209;			
Matches 940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	40	ATGGGCGCGCGGGCGCTGCTGTGGCGCTGCTGCTGCTGGCTGGAGTTCAGGAAG	99
Db	1	ATGGGCGCGCGGGCGCTGCTGTGGCGCTGCTGCTGCTGGCTGGAGTTCAGGAAG	60
QY	100	CCGGAGTCGAGGAGCGCGCGCTTATCAGGACCATGGCGCGGTCATCACCTCG	159
Db	61	CCGGAGTCGAGGAGCGCGCGCTTATCAGGACCATGGCGCGGTCATCACCTCG	120
QY	160	CGCATCGTGGTGAGAGGACGCCGAATCTGGCGCTGTGGCGTGGAGGGAGCTGC	219
Db	121	CGCATCGTGGTGAGAGGACGCCGAATCTGGCGCTGTGGCGTGGAGGGAGCTGC	180
QY	220	CTGTGGATTCGACGTATCGGAGTAGCCCTGCTACGCCACCGCTGGSCACTAC	279
Db	181	CTGTGGATTCGACGTATCGGAGTAGCCCTGCTACGCCACCGCTGGSCACTAC	240
QY	280	GCGCACTGCTTTGAACCTATAGTACCTTAGTGATCCCTCCGGGTGATGTCAGTT	339
Db	241	GCGCACTGCTTTGAACCTATAGTACCTTAGTGATCCCTCCGGGTGATGTCAGTT	300
QY	340	GGCCAGCTGACTTCATGCCATCTCTTGAGCCCTGAGGCTACTACACCGTTACTTC	399



```
Db 901 CTCCTTTTCCCTCTCTCTGCGGCTCTCCCACTCCTCGGGCGCGGTC 945
RESULT 14
LOCUS AX321357
DEFINITION Sequence 374 from Patent WO0177169.
ACCESSION AX321357
VERSION AX321357.1 GI:17905203
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lodes,M.J., Wang,T., Mohamath,R. and Indirias,C.Y.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0177168-A 374 18-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
BASE COUNT 90 a 161 c 162 g 115 t 1 others
ORIGIN
Query Match 45.6%; Score 493.4; DB 6; Length 529;
Best Local Similarity 98.6%; Pred. No. 1e-104;
Matches 510; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
QY 45 CGCGCGGGGCGCTGCTGCTGGGCGCTGCTGCTGGCTGGGCTGAGTCAAGAACCGGA 104
Db 19 CGCGCGGGGCGCTGCTGCTGGGCGCTGCTGCTGGCTGGGCTGAGTCAAGAACCGGA 78
QY 105 GTCGAGGAGGCGGCGCTTATCAGGACCATCGGCGGCGGCGGCTATCAGTGGCGCAT 164
Db 79 GTCGAGGAGGCGGCGCTTATCAGGACCATCGGCGGCGGCGGCTATCAGTGGCGCAT 138
QY 165 GTGGGGTGGAGAGACGCCGAACCTCGGGCGCTGGCGGCTGGCAGGGAGCCTCGGCTGTG 224
Db 139 CGTGGTGGAGAGACGCCGAACCTCGGGCGCTGGCGGCTGGCAGGGAGCCTCGGCTGTG 198
QY 225 GGATCCACGATGCGGAGTGAAGCTGCTCAGCCACCGCTGGGCACTCAAGCGGCGCA 284
Db 199 GGATCCACGATGCGGAGTGAAGCTGCTCAGCCACCGCTGGGCACTCAAGCGGCGCA 258
QY 285 CTGCTTTGAACCTATAGTGACCTTAGTATCCCTCGGGGTGATGGTCCAGTTGGCCA 344
Db 259 CTGCTTTGAAC-----TGACCTTAGTATCCCTCGGGGTGATGGTCCAGTTGGCCA 312
QY 345 GCTGACTTCCATGCCATCTCTTGGAGCCTGAGGCTACTACACCGCTACTTCTGATC 404
Db 313 GCTGACTTCCATGCCATCTCTTGGAGCCTGAGGCTACTACACCGCTACTTCTGATC 372
QY 405 GAATATCTATCGAGCCCTCGTACCTGGGGATTCACCTATGACATTCCTTGGTGA 464
Db 373 GAATATCTATCGAGCCCTCGTACCTGGGGATTCACCTATGACATTCCTTGGTGA 432
QY 465 GCTGTCTGCACCTGTCACTACCTAAACACATCCAGCCCATCTGCTCCAGGCTCCAC 524
Db 433 GCTGTCTGCACCTGTCACTACCTAAACACATCCAGCCCATCTGCTCCAGGCTCCAC 492
QY 525 ATTGAGTTGAGACCGGACAGAGTGTGGGTGACT 561
Db 493 ATTTGAGTTGAGACCGGACAGAGTGTGGGTGACT 529
RESULT 15
LOCUS AY005145
DEFINITION Mus musculus testisin mRNA, complete cds.
ACCESSION AY005145
```

```
AY005145.1 GI:13470305
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1055)
AUTHORS Scarman,A.L., Hooper,J.D., Boucaut,K.J., Sit,M.L., Webb,G.C.,
Normyle,J.F. and Antalis,T.M.
TITLE Organization and chromosomal localization of the murine Testisin
gene encoding a serine protease temporally expressed during
spermatogenesis
JOURNAL Eur. J. Biochem. 268 (5), 1250-1258 (2001)
MEDLINE 21153229
PUBMED 11231276
REFERENCE 2 (bases 1 to 1055)
AUTHORS Scarman,A.L., Hooper,J.D., Webb,G.C., Normyle,J.F., Sit,M.L. and
Antalis,T.M.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2000) Cellular Oncology Laboratory, Queensland
Institute of Medical Research, 300 Herston Road, Brisbane, Qld
4006, Australia
FEATURES
source Location/Qualifiers
BASE COUNT 228 a 292 c 282 g 253 t
ORIGIN
Query Match 45.2%; Score 489; DB 10; Length 1055;
Best Local Similarity 70.3%; Pred. No. 1.1e-103;
Matches 692; Conservative 0; Mismatches 275; Indels 18; Gaps 2;
QY 82 CGGCGTGGACTCAGGAAGCGGAGTCGAGGAGCGGCGGCTTATCAGGACCATCGGC 141
Db 82 CAGGTGGATCCCGAGAAACCGAACTGCAGAACCGGACTATTGTGAGGCGCTGCGGT 141
QY 142 CGACGGGTCTATCAGTCGCGCATCGTGGTGGAGAGAGCGGCAACTCGGCGCTGGCCG 201
Db 142 CACAGGACCATCCCTTCCGCTATAGTGGGTGGGTGGGTGATGCTGAGCTTGGCGCTGGCCG 201
QY 202 TGGCAGGGAGCGCTGCGCTGTGGGATTCGCCAGTATCGGAGTGAAGCTATAGTGACCTGCTCAGCCAC 261
Db 202 TGGCAAGGAGCGCTGCGTGTATGGGCAACCACTTATGTGGCGCAACCTTGTCTCAACCGC 261
QY 262 CGCTGGGCACTCAGCGCGGCGCATCTGTTTGAACCTATAGTGACCTTGTAGTATCCCTCC 321
Db 262 CGCTGGGTCTTACAGCTGCCCATCTGCTTCCAAA-----AGGATAACGATCTCTTT 312
QY 322 GGGTGGATGGTCCAGTTTGGCAGCTGACTTCCATGCCATCTCTTGGAGCGCTGAGCGC 381
Db 313 GACTGGACAGTCCAGTTTGGTGGCTGACTTCCAGGCCATCTCTCTGGAACCTACAGGCC 372
QY 382 TACTACACCGCTTACTTGGTATCGAATATCTATCTAGCCCTCGCTGCTGGGAATTC 441
Db 373 TATTCACACCGCTTACCAATAGAGATATTTCTCTGAGCCCAAGTACTCGGAGCAT 432
QY 442 CCCTATGACATTTGCTTGGTGAAGCTGCTCGACCTGTCTACCTACACTAAACATCCAG 501
```

```

Db 433 CCCAATGACATAGCCCTGCTGAAGTGTGCATCTCCAGTCACCTACAATAACTTCATCCAG 492
QY 502 CCCATCTGCTCCAGGCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGGTGACT 561
Db 493 CCCATCTGCTCCAGGCTCCACATTTGAGTTGAGAACCGGACAGACTGCTGGGTGACC 552
QY 562 GGCTGGGGGTACATCAAGAGGATGAGGCACTGGCCATCTCCCCACACACCTCCAGGAAGTT 621
Db 553 GGCTGGGGGGCTATTGGAGAGATGAGAGTCTGCCATCTCCACACACTCTCCAGGAAGTG 612
QY 622 CAGTGGCCATCAATAACACTCTATGTGCAACACACTCTTCCTCAAGTACAGTTTCGGC 681
Db 613 CAGGTAGCTATTATCAACAACAGCATGTGTAAACCATATGTACAAAAAGCCAGACTTCGGC 672
QY 682 AAGGACATCTTTGAGAGACATGTTTGTCTGGCAATGCCCAAGGGGGAAGGATGCCCTGC 741
Db 673 ACGAACATCTGGGAGAGACATGTTTGGCTGGCACCTCTGTAAGGTGGCAAGGATGCCCTGC 732
QY 742 TTCGGTGACTCAGGTGGACCCCTTGCCCTGTAAACAAGATGGACTGTGGTATCAGATTGGA 801
Db 733 TTTGGTGACTCGGAGGAGCCCTTGCCCTGGCAGCAGGATACGGTGTGGTATCAGGTGGA 792
QY 802 GTCGTGAGCTGGGAGTGGGCTGTGGTGGGCCCAATCGGGCCGGTGTCTACACCAATATC 861
Db 793 GTTGTGAGCTGGGGAATAGGCTGTGGTGGCCCAATCGCCCTGGAGTCTATACCAACATC 852
QY 862 AGCCACCACATTTGAGTGGATCCAGAAGCTGATGGCCAGAGTGGCATGTCCAGCCAGAC 921
Db 853 AGTCATCACTACAACCTGGATCCAGTCAACCATGATCCGAATGGGCTGCTCAGGCCCTGAC 912
QY 922 CCCCTCTGGCCACTACTCTTTTCCCTCTCTCTGGGCTCTCCCACTCTCTGGGGCCGGTC 981
Db 913 CCAGTCCCTCTGGCTACTGTTTCTTACTCTGGCCTGGGCTTCTCTTGGTGGAGCCCTGCC 972
QY 982 TGAGCCCTACCTGAGCCCATGAGCCCTGGGGCCACTGCCAAGTCAAGGCCCTGGTCTCTTC 1041
Db 973 TGAGCCCAACAGGTACGTACACCTGTG-----AGGTCAGGGTGTGTCTCTTTTG 1023
QY 1042 TGTCTTGTGTTGGTAATAAACACATT 1066
Db 1024 TATCTTGTGCTAATAAACCTGTT 1048

```

Search completed: May 10, 2003, 04:20:48  
Job time : 3364 secs

**THIS PAGE BLANK (USPTO)**



09/787, 844

SYSTEM:OS - DIALOG OneSearch

File 5:Biosis Previews(R) 1969-2003/Oct W1

(c) 2003 BIOSIS

File 73:EMBASE 1974-2003/Oct W1

(c) 2003 Elsevier Science B.V.

File 155:MEDLINE(R) 1966-2003/Oct W1

(c) format only 2003 The Dialog Corp.

**\*File 155: Medline has been reloaded and accession numbers have changed. Please see HELP NEWS 155.**

File 159:Cancerlit 1975-2002/Oct

(c) format only 2002 Dialog Corporation

**\*File 159: Cancerlit ceases updating with immediate effect.**

Please see HELP NEWS.

Set	Items	Description
S1	8959	CGS
S2	588	S1 AND CANCER?
S3	51	S2 AND (GYNECOL? OR TESTI?)
S4	44	RD (unique items)

5/9/1 (Item 1 from file: 5)

DIALOG(R)File 5:Biosis Previews(R)

(c) 2003 BIOSIS. All rts. reserv.

11134686 BIOSIS NO.: 199799755831

**A novel in vitro and in vivo breast cancer model for testing inhibitors of estrogen biosynthesis and its action using mammary tumor cells with an activated int-5/aromatase gene.**

AUTHOR: Tekmal Rajeshwar R(a); Durgam Vijayender R

AUTHOR ADDRESS: (a)Dep. Gynecol. Obstet., Winship Cancer Cent., Emory Univ. Sch. Med., 4219 WMB, 1639 Pierce Drive,\*\*USA

JOURNAL: Cancer Letters 118 (1):p21-28 1997

ISSN: 0304-3835

RECORD TYPE: Abstract

LANGUAGE: English

ABSTRACT: We recently showed that the cellular gene int-5/aromatase in BALB/c mammary alveolar hyperplastic nodule (D2 HAN/D2 tumor cells) is activated as a result of mouse mammary tumor virus integration within the 3' untranslated region of the aromatase gene. In the present study, we evaluated the effect of various aromatase inhibitors on androstenedione-mediated tumor cell growth. Also, we compared the effect of the non-steroidal aromatase inhibitor (CGS 16949A) on the inhibition of tumor growth. Our results show that D2 tumor cells respond well to various aromatase inhibitors and antiestrogens. We examined the usefulness of this model by using D2 tumor cells to simulate postmenopausal breast cancer employing both in vitro cell culture and in vivo ovariectomized (OVX) nude mouse. Unlike DMBA-induced tumors or other models, D2 tumor cells form very rapid tumors within a few days in intact mice or OVX nude mice with androstenedione supplementation and respond well to an aromatase inhibitor. This model with its known mechanism of aromatase activation should be useful for studying the role of intra-tumoral estrogen in mammary cancer, for evaluating the effects of aromatase inhibitors and antiestrogens, and for comparing breast cancer treatments.

5/9/2 (Item 2 from file: 5)

DIALOG(R)File 5:Biosis Previews(R)

(c) 2003 BIOSIS. All rts. reserv.

09457072 BIOSIS NO.: 199497465442

**Aromatase inhibitors - mechanisms for non-steroidal inhibitors.**

AUTHOR: Vanden Bossche Hugo(a); Moereels Henri; Koymans Luc M H

AUTHOR ADDRESS: (a)Dep. Comparative Biochem., Janssen Research Foundation, B-2340 Beerse\*\*Belgium

JOURNAL: Breast Cancer Research and Treatment 30 (1):p43-55 1994

ISSN: 0167-6806

DOCUMENT TYPE: Literature Review

RECORD TYPE: Abstract

LANGUAGE: English

ABSTRACT: The conversion of androgens to estrogens occurs in a variety of

**THIS PAGE BLANK (USPTO)**

cells and tissues, such as ovarian granulosa and **testicular** cells, placenta, adipose tissue, and various sites of the brain. The extragonadal synthesis of estrogens has great pathophysiological importance. Estrogens produced by, for example, adipose tissue have a role in the pathogenesis of certain forms of breast **cancer** and endometrial adenocarcinoma. The biosynthesis of estrogens is catalyzed by the aromatase, an enzyme localized in the endoplasmic reticulum that consists of two components: a cytochrome P450 (P450 Arom, P450 19 product of the CYP19 gene) and the NADPH cytochrome P450 reductase. The alignment of the amino acid sequences of human P450 19 with other mammalian P450s shows little sequence similarity, which indicates not only that P450 19 is a unique form of the P450 superfamily but also that the aromatase may be a good target, for the development of selective P450 inhibitors. Aminoglutethimide (AG) is the pioneer drug of the reversible competitive nonsteroidal aromatase inhibitors. Since AG is a nonspecific aromatase inhibitor and presents some problems with tolerability, a number of structural analogues have been synthesized. For example, rogletimide is slightly less potent than AG but has the advantage of not inhibiting the cholesterol side-chain cleavage and is devoid of sedative action. Elongation of the ethyl substituent of AG and rogletimide leads to an increase in aromatase inhibition. Further studies led to the discovery of a new generation of much more potent aromatase inhibitors. An example is fadrozole. However, although fadrozole is a poor inhibitor of the cholesterol side-chain cleavage, it suppresses aldosterone release by ACTH-stimulated human adrenocortical cells. More selective aromatase inhibitors are the triazole derivatives. Examples are **CGS** 20267, **CGS** 47645, R 76 713, and ICI D1033. R 76 713's aromatase inhibitory effect is largely due to its (+)-S-enantiomer, vorozole. Computer modeling studies of the interaction of vorozole with part of the "I-helix" of P450 19 suggest that the chlorine-substituted phenyl ring of vorozole interacts with the gamma-carbonyl group of Glu-302. Thr-310, which corresponds to the highly conserved Thr-252 in P450 101, interacts with vorozole's triazole ring, and the 1-methyl-benzotriazole moiety binds near Asp-309.

REGISTRY NUMBERS: 9039-48-9: AROMATASE; 112809-51-5: **CGS** 20267;  
118949-22-7: R 76713; 129731-10-8: VOROZOLE; 121840-95-7: ROGLETIMIDE

**THIS PAGE BLANK (USPTO)**

*Search notes***WEST****The Contents of Case 09787844**

---

Qnum	Query	DB Name	Thesaurus	Operator	Plural
Q1	cancer specific antigen	USPT	ASSIGNEE	ADJ	YES
Q2	Q1 and ovarian cancer	USPT	ASSIGNEE	ADJ	YES
Q3	cancer specific antigen.clm	USPT	ASSIGNEE	ADJ	YES
Q4	her2.clm	USPT	ASSIGNEE	ADJ	YES
Q5	her2.clm.	USPT	ASSIGNEE	ADJ	YES
Q6	cancer specific antigen.clm.	USPT	ASSIGNEE	ADJ	YES

---

[Run Case](#)[Update Case](#)[Cancel](#)[Help](#)[Main Menu](#)[Logout](#)

**THIS PAGE BLANK (USPTO)**